

Accession	Sequence	Position
Dp	GGAGCTGTGATGAATTTGGGTGAGGAGGCTTAACATCAGATGTGACGTGCAAAATCTATGAT	376
Qy	TCCCTGCTGGCTCTTTCTTCGAACTTAAGGACGACAGAGACTGATGTGTGCTGCTCC	360
Dp	TCCCTGCTGGCTCTTTCTTCGAACTTAAGGACGACAGAGACTGATGTGTGCTGCTCC	436
Qy	GTGATGTCTCTCTTGGCTTTCATGATGGCCATCCTTGGCATGAATCACCAGGTGCACG	420
Dp	GTGATGTCTCTCTTGGCTTTCATGATGGCCATCCTTGGCATGAATCACCAGGTGCACG	495
Qy	GGGACATAGAAAGGTGAAGGCTCACTTCTGTGACGCTGGAAATCATCTTCATATC	480
Dp	GGGACATAGAAAGGTGAAGGCTCACTTCTGTGACGCTGGAAATCATCTTCATATC	555
Qy	ACGGGATGATGTGCTCATCCCTGTGAAGTGGGTGGCAATGCATCATCAGAGATT	538
Dp	ACGGGATGATGTGCTCATCCCTGTGAAGTGGGTGGCAATGCATCATCAGAGATT	610

RESULT 2
US-09-620-312D-986
Sequence 986, Application US/09620312D

```

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wenman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhilwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radost T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784C1P28
CURRENT APPLICATION NUMBER: US/09/620.312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_Fl_genes Version 1.0
SEQ ID NO 986
LENGTH: 708
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (46)..(708)
IS-09-620-312D-986

```

Query Match	8.7%;	Score 175.8;	DB 4;	Length 708;
Best Local Similarity	57.4%;	Pred. No. 3.8e-36;		
Matches 337;	Conservative	0;	Mismatches 247;	Indels 3; Gaps 1

QY 6 TCACAGTGGCTAAATCATCCACAGATATATGGCAACCCATCCTTAGAAATCGGTG 125
 15 TCCTTGCACTGACAGCTCTTCAACCTCGCCATGCGCTCTCGCGAATGACAGATCTCGG 74
 126 GCTGTTCTTTGAGTGTGTAATGTGTGGACACAGTGTGTCACTGTCACTGCTCAATG 185
 75 AGTCGTCCTGACACTGCTGAGCTGGATGTAATGCTGTATCTCTGTGCTCCGCGCCATG 134
 Db

OY	186	GAGAGTGTGGGCTTCATTTGAAAAACAATGTTGTTTAAAAACTCTGGAGAGACT	245
Db	135	GAGGTGACCGCTTTCATTCGGCAACAGCATCGTGGTGCCAGGTGGTGGAGGCGCT	194
OY	246	GTGGATGAATTGGCTGTGAGGCGAGGCTTAACATCAGATGTGAGTGCAAAACTCATGATTTCCCT	305
Db	195	GTGGAATGTCCTCGTGGTGCACAGACACCGGCGAATGAGTGAAGGTATACGACTACT	254
OY	306	GCTGGCTCTTTCTCCGGAACCTACAGAGCAGCAGAGAGATGTATGTGCTGCTTCCGAT	365
Db	255	GCTGGGCGCTGCCACAGAACTTGAGGCTGCAGTGTGCCCTCTGTGTCAATCCGCCCTCTGT	314
OY	366	GTCTCTTTGGCTTTCATGATGATGAGCCATCTTGGCATGAATGCAACAGTGCACGCGGGA	425
Db	315	GGCGCTGTTCGCGCTTGCTGATCTTACCTTGCTGGGCGCAAGTGTCCACCTGTGTGGAGA	374
OY	426	CAATGAGAGTGGAGAGGCTCACTTCGTGTGACGSGCTGGAAATCATCTTCAATCAACAGG	485
Db	375	GA--AGGATTCCAAGGCGCGCGCTGATCTACCTTGGATTGTCTTTGTCACTCAGG	431
OY	486	CATGTGTGCTCATCCCTGTGAGCTGGATTTGCCATGCCATCATCAGAGATTTCTATA	545
Db	432	GGTCTGACGCTAATCCCGGTGTGTGTGAGACGGCGGATCCGTATCCGGGACTCTTATTA	491
OY	546	CTCAATAGTGAATGTTTGCCAAAAACGTGAGCTTGGAGAGCTCTTACTTAGATGAC	605
Db	492	CCCCCTGGTGGCTGAGGCGCCAAAACGGGAGCTGGGGGCTCCCTCTAATCTGGGCTGGC	551
OY	606	CACGGACTGTGCTGATGTTGTGGAGAGCTCTGTCTGCTGCCTTT	652
Db	552	GGCTCAGGCTTTTGTGTCTGGGTGGGGGTGTGTCGTGCACCTT	598

RESULT 3
US-09-205-258-106
; Sequence 106, Application US/09205258

```

1  GENERAL INFORMATION:
2  APLICANT: Young et al.
3  TITLE OF INVENTION: 207 Human Secreted Proteins
4  FILE REFERENCE: P2007P1
5  CURRENT APPLICATION NUMBER: US/05/205,258
6  CURRENT FILING DATE: 1998-12-04
7  EARLIER APPLICATION NUMBER: PCT/US98/11422
8  EARLIER FILING DATE: 1998-06-04
9  EARLIER APPLICATION NUMBER: 60/048,885
10 EARLIER FILING DATE: 1997-06-06
11 EARLIER APPLICATION NUMBER: 60/049,375
12 EARLIER FILING DATE: 1997-06-06
13 EARLIER APPLICATION NUMBER: 60/048,881
14 EARLIER FILING DATE: 1997-06-06
15 EARLIER APPLICATION NUMBER: 60/048,880
16 EARLIER FILING DATE: 1997-06-06
17 EARLIER APPLICATION NUMBER: 60/048,896
18 EARLIER FILING DATE: 1997-06-06
19 EARLIER APPLICATION NUMBER: 60/049,020
20 EARLIER FILING DATE: 1997-06-06
21 EARLIER APPLICATION NUMBER: 60/048,876
22 EARLIER FILING DATE: 1997-06-06
23 EARLIER APPLICATION NUMBER: 60/048,895
24 EARLIER FILING DATE: 1997-06-06
25 EARLIER APPLICATION NUMBER: 60/048,884
26 EARLIER FILING DATE: 1997-06-06
27 EARLIER APPLICATION NUMBER: 60/048,894
28 EARLIER FILING DATE: 1997-06-06
29 EARLIER APPLICATION NUMBER: 60/048,971
30 EARLIER FILING DATE: 1997-06-06
31 EARLIER APPLICATION NUMBER: 60/048,964
32 EARLIER FILING DATE: 1997-06-06
33 EARLIER APPLICATION NUMBER: 60/048,882
34 EARLIER FILING DATE: 1997-06-06
35 EARLIER APPLICATION NUMBER: 60/048,899
36 EARLIER FILING DATE: 1997-06-06

```

```

; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06

```

```

/ EARLIER APPLICATION NUMBER: 60/048, 893
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 900
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 901
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 892
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 915
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049, 019
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 970
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 972
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 916
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049, 373
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 875
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049, 374
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 917
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 949
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 974
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 883
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 897
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 898
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 962
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 963
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 877
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048, 878
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/070, 923
/ EARLIER FILING DATE: 1997-12-18
/ EARLIER APPLICATION NUMBER: 60/092, 921
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/094, 657
/ EARLIER FILING DATE: 1998-07-30
/ NUMBER OF SEQ ID NOS: 1227
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 106
/ LENGTH: 1705
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (724)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-203-258-106

```

```

Query Match      7.9%; Score 159.2; DB 4; Length 1705;
Best Local Similarity 56.6%; Pred. No. 1.3e-31;
Matches 315; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

```

```

QY 92 GGATATGGCAACCATGCTTGAAGATCGCTGGGCTTTCTTGCTGCTGTTGAATGG 151
DB 204 GAACATAGGCTCCATGAGGCTACAGATATGGGCTGCGCTGCGCTGCGCTGCG 263
QY 152 TGGGACAGTGGCTGTCACTGTCACTGCTGCTGAGTGTGAGAGTGTGAGTGTG 211
DB 264 TGGCGGTATGCTGTGCTGTGCGGCTGCGGCTGCCCATGTGAGCGGCTTATCGG 323

```

```

QY 212 ACATGCTGTTTTTGAAGAACTTCTGGAGAGAGTGTGATGAATGGCTGAGCAGGCTA 271
DB 324 ACATTGTACCTCGCAGAGCCATCTGGAGAGGCTTATGATGAACTGTGGTGTGACAGCA 383
QY 272 ACATGAGATGCAAGTGCAGAAATCTATGATTCCTGCTGCTGCTTTCTCCGACCTTACAG 331
DB 384 CCGGCAATGCAAGTGCAGAAAGTGTACACTGCTGCTGCGCACTGCCAGAGACTGTCAGG 443
QY 332 CAGCCAGAGAGCTGATGTGTGCTGCTCCGATAGTCTTCTGCTTATGATGATGAGCA 391
DB 444 CGGCGCGCGCTGCTGCTATCATCATGAGCATCATGCTGCTGCTGCTGCTGCTGCTGCTG 503
QY 392 TCCTTGGCATGAATGCAACGAGTGCACGGGGCAATGAGAGTGAAGGCTTACATTTC 451
DB 504 TGGTGGGGGCAAGTGTACCACTGCTGGAGG---ATGAAGGGCCAGAGCCAGACCA 560
QY 452 TGCTGACGGCTGAGATCATCTTTCATCATACGGGCATGAGTGTGCTCATCCGTGAGCT 511
DB 561 TGATGTGGCGGGCTGTGTCTGCTGTGGCGGCTTATGATGATGCTGCTGCTGCT 620
QY 512 GGGTTGCCAATGCAATCATCATGAGATTTCTAATACTCAATGATGATGTTGCCCAAAAC 571
DB 621 GGAAGGCGCCCAATCATATCATCAAGACTTCTAATACTCGGTGTGCTGCTGCGGAGAGC 680
QY 572 GTGAGCTTGGAGAAAGCTCTCTTACTTATGATGAGCAACAGGCACTGTGCTGATTTGGAG 631
DB 681 GGAAGATGGGTGCTGCTGCTTACGTCGCTGCGGCGGCTGCGGCTGCTGCTGCTGCG 740
QY 632 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
DB 741 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757

```

```

RESULT 4
US-09-300-958A-22
/ Sequence 22, Application US/09300958A
/ Patent No. 6495319
/ GENERAL INFORMATION:
/ APPLICANT: McCelland, Michael
/ APPLICANT: Welsh, John
/ TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of
/ FILE REFERENCE: P-PH 3457
/ CURRENT APPLICATION NUMBER: US/09/300, 958A
/ CURRENT FILING DATE: 1999-04-27
/ PRIOR APPLICATION NUMBER: 60/083, 331
/ PRIOR FILING DATE: 1998-04-27
/ PRIOR APPLICATION NUMBER: 60/098, 070
/ PRIOR FILING DATE: 1998-08-27
/ PRIOR APPLICATION NUMBER: 60/118, 624
/ PRIOR FILING DATE: 1999-02-04
/ NUMBER OF SEQ ID NOS: 85
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 22
/ LENGTH: 1665
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-300-958A-22

```

```

Query Match      7.9%; Score 158.6; DB 4; Length 1665;
Best Local Similarity 56.6%; Pred. No. 1.8e-31;
Matches 315; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

```

```

QY 92 GGATATGGCAACCATGCTTGAAGATCGCTGGGCTTTCTTGCTGCTGTTGAATGG 151
DB 178 GAACATAGGCTCCATGAGGCTACAGATATGGGCTGCGCTGCGCTGCGCTGCG 237
QY 152 TGGGACAGTGGCTGTCACTGTCACTGCTGCTGAGTGTGAGAGTGTGAGTGTG 211
DB 238 TGGCGGTATGCTGTGCTGTGCGGCTGCGGCTGCCATGTGTGAGCGGCTTATCGG 297

```

QY 212 ACATCGTGTGTTTGAACCTTCTGGAGAGAGTGTGATGAATTGCGTAGGACAGCTTA 271
| | | | |
Db 298 ACATTGTCACTCGCAGACATCTGGAGAGGCGCATGATGAATCGTGATGTCAAGACA 357
| | | | |
QY 272 ACATAGAGATGATGACAAATCTATGATTCCTGCTGCTCTTTCTCCGAGCTTCAAG 331
| | | | |
Db 358 CCGGCAATGATGACAGAGGTAGCACTGCTGCTGCGACATGCGGACAGACTTGCAG 417
| | | | |
QY 332 CAGCCAGAGAGTATGATGCTGCTTCCGATGATGCTCTTCTGCTTTCATGATGAGCA 391
| | | | |
Db 418 CCGCCCGGCGCTCGTATCATCATGACATCATCGGCTGCTGCTGCGGATGCTGCTCCG 477
| | | | |
QY 392 TCCTTGGATGAATGACCAAGGTCACAGGAGCAATGAGAGGTGAAGGCTCAATTC 451
| | | | |
Db 478 TGTGAGGAGGAGAGTGTACCAAGTCTGAGAG--ATGAAAGCGCCAGGCGCAAGACA 534
| | | | |
QY 452 TGCTAGCGCTGATGATCATCTTCAATCAAGGAGTGTGCTGATCCCTGAGCT 511
| | | | |
Db 535 TGATGCTGCGGCGGTGTCTTCTGTTGCGCGCTTATGATGATGCGGTCTCT 594
| | | | |
QY 512 GGGTGGCAATGATCATCATGAGATTTCTATACTCAATGATGAATGTTGCCAAAG 571
| | | | |
Db 595 GAGACGCGCCCAAGACATCATCAAGATTTCAATCCGCTGATGCTCCGCGCAGAAC 654
| | | | |
QY 572 GTGAGCTTGGAGAGTCTCTACTTAGATGAGACAGGACATGATGATGTTGAG 631
| | | | |
Db 655 GAGAGATGAGGCTCTCTTACGTCGCTGAGGCGCTCCGCGCTGCTCTTGGCG 714
| | | | |
QY 632 GAGCTGTGCTGTCTGCTGC 648
| | | | |
Db 715 GGGGCTGCTTGTGCTGC 731
| | | | |

RESULT 5
US-09-673-395A-71
; Sequence 71, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALERE-12
; CURRENT APPLICATION NUMBER: US/09/673.395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-71

Query Match 7.5%; Score 150; DB 4; Length 1253;
Best Local Similarity 55.1%; Pred. No. 2.7e-29;
Matches 314; Conservative 0; Mismatches 255; Indels 1; Gaps 1;
QY 115 GAAATCGCTGGCTGTTTCTTGTGTGTGTGAT-GETGGGACAGTGGCTGTCACTG 173
| | | | |
Db 363 GCAAGTTCGGGCTTCTCCATGCGCCCTGCTGGGTGAGTGTGTGTGTGTGACCCG 422
| | | | |
QY 174 CATGCTCAGTGGAGAGTGTGCTTCAATGAAAAACAATCGTGTGTTTGAACCTT 233
| | | | |
Db 423 CATCCCGAGATGCGAGATGAGCTCTATGCGGAGACATATATACGCGCCAGCCAT 482
| | | | |
QY 234 CTGGAGAGAGTGTGATGAATTGCTGAGGACGCTTACATCAGATGACATGCAAAAT 293
| | | | |
Db 483 GTACAAAGGAGGCTGTGATGATGACTGCTGACCAAGACAGGAGATGATGACAAAT 542
| | | | |
QY 294 CATATATCCCTGCTGCTGCTTCTTCCGAGACTACAGGACAGGACATGATGTGTG 353
| | | | |

Db 543 GTACGATCGGTGCTCGCCCTGTCCGCGCTTGACAGGCTCATCGACCTTATGTGTGT 602
| | | | |
QY 354 TGCTTCGATGATCTCTTCTTGGCTTTCATGATGCGCATCTTGGCATGAAATGACAG 413
| | | | |
Db 603 CTCCTGTGCTGTGGCTTCTGTGCGCATGTTTGTGCGAAGATGGCATGAGATGACCG 662
| | | | |
QY 414 GTGACGCGGAGCAATGAGAGGTGAAGGCTCACTTCTGCTGACGCTGGAATATCTT 473
| | | | |
Db 663 CTGTGGGGAGACGAAGATGAGAGGCGCGTATAGCAAGGATGAGAGGATATTTT 722
| | | | |
QY 474 CATATCAGGAGGATGATGTGTCTCATCCCTGTGAGCTGGTGTGCAATGACCATCATG 533
| | | | |
Db 723 CATGTGAGAGTCTTGTGCGCTTGTGATGCTTCTCTGTATGCGCATCATGATGTGAC 782
| | | | |
QY 783 AGACTTTATTAACCTTTGATCCCTACCAATTAATGATGATGATGATGATGATGAT 842
| | | | |
Db 594 CTATGATGACCAAGCAGTGTGCTGATGATGAGAGGCTGATGATGATGATGATGATG 653
| | | | |
QY 843 TATGAGCTGAGAGGATGCTGCTTCTGATCATCTGAGAGTGCATCTGCTCTCTCTG 902
| | | | |
QY 654 TTGTTGCAAGAAAGAGAGTGTGACAG 683
| | | | |
Db 903 TCTGAGATGAGAGAGGCTGATACCG 932
| | | | |

RESULT 6
US-09-489-847-93
; Sequence 93, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489.847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 93
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-93

Query Match 7.4%; Score 149; DB 4; Length 1722;
Best Local Similarity 55.3%; Pred. No. 5.8e-29;
Matches 306; Conservative 2; Mismatches 242; Indels 3; Gaps 1;
QY 97 ATGCAACCATGCTTAAATATCGCTGGCTGTTCTTGTGTGTGTGTGATGATGATG 156
| | | | |
Db 461 ATGGGATCCCAAGCTTGAATCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 520
| | | | |
QY 157 AAGTGTCTGATGCTGCTCAATGAGAGTGTGAGGCTTCAATGAAAAACAATC 216
| | | | |
Db 521 CTGATCTGCGGCTGCGGCTGCCCATGTGCAAGTACCGCTTCTCTGACCAACAATC 580
| | | | |
QY 217 GTGATTTTGAACCTTCTGGAAGAGCTGTGATGAATGCTGTGAGGAGGCTTAATC 276
| | | | |
Db 581 GTGACGCGGAGACCACTGGAAGGAGGCTGTGATGTCTGTGTGTGTGTGTGTGTGTG 640
| | | | |

QY 277 AGATGACATGCAAAATCTGATGATTCCTGCTGGCTTTCTCCGACCTACAGGACGCC 336
DB 641 CACATGACATGCAAAAGTGTACGACTGGGTGCTGGCTGACACCCGAGGTGACAGCGCGC 700
QY 337 AAGAGACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGCTTCATGATGGCCATCTT 356
DB 701 CCGGCGCTCACCGTGTAGCGCGCTGCTGCTGGGTGCTTGCGCTTCTGTAACCTTGCGC 760
QY 397 GGCATGAAATGCAACGAGTGTGCAAGGAGGACAAATGAAAGGTGAAGGTCAATTCCTG 456
DB 761 GCGCGGACATGCAACCACTGCG--GTGGCCCCGGGCGGCAAGCGCGGTGTGGCCTC 817
QY 457 AGGCTGGAATCATCTTATCATCAAGGACATGTGTGCTCATCTCTGTGATGAGTTG 516
DB 818 ACGGAGAGGCTGCTTACTCTGTTTTCGCGGCTGCTGGCGCTGTGCACTCTGCTGTTTC 877
QY 517 GCCAATGCCATCATGAGATTTCTATACTCAATAGTGAATGTTGCCAAAAACGTGAG 576
DB 878 GCCAATGTTGTGCTCGCGAGTTTACGACCCCGTCTGTGCCCGTGTGCGAAGATAGAG 937
QY 577 CTTGAGAACTCTCTACTTAGATGACACAGGACATGTGTGATTTTGAAGAGCT 636
DB 938 CTGGGCGAGCGCTGTACATCGGCTGGGCGGCGCACCGGCTGTGATGTAGGCGGCTGC 997
QY 637 CTGTTCTGCTGCG 649
DB 998 CTCTTGTGCTGCG 1010

RESULT 7

US-09-886-683A-3
Sequence 3, Application US/09886683A
Patent No. 6627439
GENERAL INFORMATION:
APPLICANT: Hoevel, Thorsten
APPLICANT: Koch, Stefan
APPLICANT: Kuddies, Manfred
APPLICANT: Mundigl, Oiaf
APPLICANT: Rueger, Petra
TITLE OF INVENTION: Antibodies against SEMP1 (p23)
FILE REFERENCE: Case 20692
CURRENT APPLICATION NUMBER: US/09/886,683A
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: EP0011344.6
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: EP01107799.7
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 3443
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (221)..(853)
US-09-886-683A-3

Query Match 6.7%; Score 134.8; DB 4; Length 3443;
Best Local Similarity 52.5%; Pred. No. 4.2e-25;
Matches 225; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 86 CCCAAGATATAGCAACCCATGCTTAGAAATGCTGGGCTGTTCTTGATGATGTTG 145
DB 210 CCGAGGAGTATGAGCAACCGGCGGCTGACGCTGTGGCTTATCTCGCTTCTCG 269
QY 146 GATGTGTGGGACAGAGGTGTCACTGTCAATGCCCAATGAGAGTGTGCGGCTTATTG 205
DB 270 GATGATCGGCGCATCTGTACGACCTGTGCCCCCAATGAGATTTACTCTTATGCCG 329
QY 206 AAAACAACATGTGTTTTTGAACCTTCTGGGAGAGACTGTGATGAATTCGTGAGGC 265

DB 330 GCGACAACATGTGACCCGCCAGGCCATGTACAGAGGAGCTGTGATGTCTCGGTGCGC 389
QY 266 AGGCTAATCATGAGATGCAAGTCAAAATCTATGATTCCTGCTGCTTTCTCCGACC 325
DB 390 AAGACACCGGAGATTCAGTGCAAGCTTTGACTCTTGCTGATCTGACGACACAT 449
QY 326 TACAGGACCCAGAGACTGATGTGTGCTGTTCCGTGATGTCTTCTTGCTTCAATGA 385
DB 450 TCGAAGCAACCGGCTGTGATGTGTGATGTCATCTCTGAGATATGCAATCTTGG 509
QY 386 TGGCATCTTGTGATGAAATGCAACGAGTGTGACAGGAGGACAAATGAGAGGTGAAGCTC 445
DB 510 TGGCACCGTTGGATGAAAGTATGAAAGTCTTGAAGACATGAGATGACAGATGA 569
QY 446 ACATTCTGTACGCTGTGAATCATCTTATCATCAAGGACATGTGTGCTCATCTCCTG 505
DB 570 GATGCTGTATGGGGGTGCAATTTCTTCTTGAGGCTGTGCTATTTTGTGCA 629
QY 506 TGAGCTGGATGCGCATGCAATGCAAGATTTCTATACTCAATAGTGAATGTTGCC 565
DB 630 CAGCATGTATGCAATGAAATCGTTCAAGATTTCTATGACCTATGACCCAGTCAATG 689
QY 566 AAAACGTGAGCTGTGAGAGACTCTCTACTTAGATGACACAGGACATGTGTGATG 625
DB 690 CAGGTACGAATTTGTGAGGCTCTTCACTGGCTGGGCTGTGCTTCTCTGCTTC 749
QY 626 TTGAGAGAGCTGTGTGCTG 647
DB 750 TGGAGGTGCTTACTTGTCTG 771

RESULT 8

US-09-130-491-3
Sequence 3, Application US/09130491
Patent No. 6416974
GENERAL INFORMATION:
APPLICANT: Holzman, Douglas A.
APPLICANT: Goodaarl, Andrew D.J.
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
FILE REFERENCE: 09404/041001
CURRENT APPLICATION NUMBER: US/09/130,491
PRIOR FILING DATE: 1998-08-07
EARLIER APPLICATION NUMBER: US 60/058,108
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: US 60/054,961
EARLIER FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 3483
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (240)...(872)
US-09-130-491-3

Query Match 6.7%; Score 134.8; DB 4; Length 3483;
Best Local Similarity 52.5%; Pred. No. 4.2e-25;
Matches 295; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 86 CCCAAGATATAGCAACCCATGCTTAGAAATGCTGGGCTGTTCTTGATGATGTTG 145
DB 229 CCGAGGAGTATGAGCAACCGGCGGCTGACGCTGTGGCTTATCTCGCTTCTCG 288
QY 146 GATGTGTGGGACAGAGGTGTCACTGTCAATGCCCAATGAGAGTGTGCGGCTTATTG 205
DB 289 GATGATCGGCGCATCTGTACGACCTGTGCCCCCAATGAGATTTACTCTTATGCCG 348
QY 206 AAAACAACATGTGTTTTTGAACCTTCTGGGAGAGACTGTGATGAATTCGTGAGGC 265
DB 349 GCGACAACATGTGACCCGCCAGGCCATGTACAGAGGAGCTGTGATGTCTGTGCTGCGC 408

```

QY 266 AGCTTACATGAGTGTGAGTGAATAATCTATGATTCCCTGCTGGCTCTTTCGGGACC 325
Db 409 AGAGCAGCGGCGAATCCAGTGAAGCTTTGCTCTCTTGCAATCTGAGCAGACAT 468
QY 326 TACAGCAGCCAGAGACTGATGTGTGCTGCTTCGGATGTCTCTTCTTGCCATTGATGA 385
Db 469 TGCAAGCAACCGGCTGCTGATGTGTGCAATCTCTCGGAGATGATGCAATCTTTG 528
QY 386 TGGCCATCTTGGCATGAATGACACAGGTGCAAGGGGAGCAATGAAAGTAAAGCTTC 445
Db 529 TGGCCACCGTTGGCATGAAGTGTGATGAAGTCTTGAAGACGATGAGGTGCAAGATGA 588
QY 446 ACATTCTGCTACAGCGCTGGAATCATCTTCATCATCAGGGGCAATGAGTGTGCTGATCC 505
Db 589 GGATGGCTGCTGATTTGGGGGTGCGAATTTCTCTTGAGGTCTGGCATTTTGTGCA 648
QY 506 TGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTTAATCTAATGATGAATGTTGCC 565
Db 649 CAGCATGTGATGCGCAATGAAATCGTTCAGAAATCTATGACCCATGACCCGATGATG 708
QY 566 AAAACGTGAGCTTGGAGAACCTCTCTACTATGAGATGACCAAGGCACTGTGCTGATTG 625
Db 709 CCAGGTACGAATTTGGTGTAGGCTCTCTTCACTGAGCTGGGCTGCTCTTCTCTGCTTC 768
QY 626 TTGAGAGAGCTCTGCTGCTG 647
Db 769 TGGAGAGTGCCTACTTGTCTG 790

```

RESULT 9

```

US-09-489-847-125
; Sequence 125, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 125
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-125

```

```

Query Match 6.4%; Score 129; DB 4; Length 1380;
Best Local Similarity 55.7%; Pred. No. 8.3e-24;
Matches 308; Conservative 0; Mismatches 240; Indels 5; Gaps 3;

```

```

QY 97 ATGGCAACCCATGCTTGAATGCTGGGCTGTTTCTGTGTGTGATGAGTGGGG 156
Db 135 ATGGGGTCCGCGACGCTGGAATCTGTGGCTGTGTGCTGTGCTGTGGGCTGGGGGGGT 194
QY 157 ACAGTGGCTGTGACATGATGCTTCACTGAGAGTGTGCGCTTTCATTTGAAAAACAATC 216
Db 195 CTGATCTCGGGGTGGGGCTGCGCATGTGCGAGGTGACCGGCTTCCCTGAGCAACAATC 254
QY 217 GTGGTTTGAATACTTCTGGAAAGACTGTGATGTAATGCGTGAAGCGAGGCTAAATC 276

```

```

Db 255 GTACGCGCGAGACCACTGGAAGGGGCTGTGATGTGTGCTGCTGATGAGACAGCGGC 314
QY 277 AGATGAGTGTGCAAAATCTATGATTCCCTGCTGGGCTCTTCTCCGGAACCTAGAGGACGC 336
Db 315 A-CATGAGAGCAAGGTATGCACTGTGTGCTGTGCTGATGACCGAGGTGCAAGCGGCG 373
QY 337 AGAGACTGATGTGTGCTGCTCCGTGATGTCTCTTGTGCTTGTATGATGAGCCATCTT 396
Db 374 CGGGCGCTACCGTGAAGCGCGGTGCTGTGGGCTTGTGGCTCTTGTGACCTGTGCG 433
QY 397 GGCATGAATGACACAGGTGCAAGGGGACATGAAAGTAAAGGTACATTTCTGTC 456
Db 434 GGCCTGAGAGTGCACACCTGCG--GTGGCCCGGGGCGCCAGAGCGGTGTGCGCCCTC 490
QY 457 ACGGCTGAATCATCTTTCATCATCAGGGCAGTGTGCTCATCTGTGATGCTGGTT 516
Db 491 ACGGAGAGCGCTGTCTTACCTGTTTTCGGGCTGTGCGCTGTGCACTGTGCTGGTT 550
QY 517 GCCAATCCATCATCAGAGATTTCTTAATCTAATGTAATGTTGCCAAAAAGTGA 576
Db 551 GCCAATCTTGTGCGCGCGAGTTTACGACCGCTGTGCTGCTGTGCAAGATGAG 610
QY 577 CTGAGAGAGCTCTTACTATGAGATGAGACCAAGGCACTGTGCTGATTGTTGAGAGAGCT 636
Db 611 CTGGGCGCA-CGCTGTACATGAGCTGGGCGGCGCACCGGCTGCTATGTAGGCGGCTGC 669
QY 637 CTGTTCTGCTGCG 649
Db 670 CTCTGTGTGCTGCG 682

```

RESULT 10

```

US-09-663-600A-45
; Sequence 45, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclercq, Aymeric
; TITLE OF INVENTION: EXTENDED CNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 45
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

FEATURE:
NAME/KEY: sig_peptide
LOCATION: 160..231
OTHER INFORMATION: Von Heljne matrix
OTHER INFORMATION: score 5.6993980926514
NAME/KEY: polyA_signal
LOCATION: 1510..1515
NAME/KEY: polyA_site
LOCATION: 1506..1519

```

```
/ NAME/KEY: misc.feature
/ LOCATION: 1048..1504
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: AA352647
/ NAME/KEY: misc.feature
/ LOCATION: 597..846
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: AA345449
/ NAME/KEY: misc.feature
/ LOCATION: 39..53
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: AA345449
/ NAME/KEY: misc.feature
/ LOCATION: 113..149
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: AA345449
/ NAME/KEY: misc.feature
/ LOCATION: 98..400
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: T86266
/ NAME/KEY: misc.feature
/ LOCATION: 1210..1489
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: T86158
/ NAME/KEY: misc.feature
/ LOCATION: 954..983
/ OTHER INFORMATION: homology
/ OTHER INFORMATION: id: AA116709
/ OTHER INFORMATION: est
US-09-663-600A-45

Query Match
Best Local Similarity 50.8%; Score 124.8; DB 4; Length 1524;
Matches 319; Conservative 2; Mismatches 304; Indels 3; Gaps 1;
```

```
577 CTGGAGAGAGCTCTCTACTAGATGATGACACAGGCACTGCTGATGTTGGAGAGCT 636
637 ATTGAGAGAGCTCTTACTTGGCACTTTCTTCCCTGCTCCCTGATAGCTGGATC 696
637 CTGTTCTGCTGCTTTTGTGCAACGAAAGACAGTACTACAGTACTGANTACT 696
697 ATCCCTGCTTTTCTGCTCATCCAGAAATGCTTCACTCACTAGATGCTTACCA 756
697 TCCCATCGACACACCAAAAAGTATC 724
757 GCCCAACTCTTGGCCACAGGAGCTCTC 784

RESULT 11
US-09-663-600A-139
/ Sequence 139, Application US/09663600A
/ Patent No. 6573068
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Malne Edwards, Jean-Baptiste
/ APPLICANT: Duclerc, Aymeric
/ APPLICANT: Bouqueleret, Lydie
/ TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
/ FILE REFERENCE: 31 US3 CIP
/ CURRENT APPLICATION NUMBER: US/09/663,600A
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/191,997
/ PRIOR FILING DATE: 1998-11-13
/ PRIOR APPLICATION NUMBER: 60/066,677
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/069,957
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/074,121
/ PRIOR FILING DATE: 1998-02-09
/ PRIOR APPLICATION NUMBER: 60/081,563
/ PRIOR FILING DATE: 1998-04-13
/ PRIOR APPLICATION NUMBER: 60/096,116
/ PRIOR FILING DATE: 1998-08-10
/ PRIOR APPLICATION NUMBER: 60/099,273
/ PRIOR FILING DATE: 1998-09-04
/ NUMBER OF SEQ ID NOS: 229
/ SOFTWARE: Patent.pm
/ SEQ ID NO 139
/ LENGTH: 1400
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 36..107
/ OTHER INFORMATION: Von Heljne matrix
/ OTHER INFORMATION: score 5.6999980926514
/ NAME/KEY: poly_a_signal
/ LOCATION: 1302..1307
/ NAME/KEY: poly_a_site
/ LOCATION: 1389..1400
US-09-663-600A-139

Query Match
Best Local Similarity 51.0%; Score 124; DB 4; Length 1400;
Matches 320; Conservative 0; Mismatches 305; Indels 3; Gaps 1;
```

```
577 ATGGCAACCATGCTTTAGAAATGCTGCTGCTTTCTTGAGTGTGGAATGCTGGC 156
637 ATGGCTCTCTTGGCTCTCACTTGTGGCTTACATCTTCTTGGCTTCTTGGCT 95
157 AAGTGGCTGTCACTGTCACTCAAGTGTGAGTGTGCTTCAATGAAACATC 216
96 AACTGTTGCTGCTGCTCTCCAGCTGGAACAAATTTATGTGCTGCTGAGCAT 155
217 GTGGTTTTGAAAACCTTGGGAAAGAGCTGTGATGATGCTGAGGAGGCTAACATC 276
156 GTGACAGCAGTTGGCTTCTCCAAAGGCTCTGTGATGATGATGATGATGATGATG 215
```

QY	27	AGGATGACAGGCAAAATCATGATATTCCTCGCTGAGCTTTTCTCGGACCTCAACAGGCGGC	336
Db	216	ATACCCAGTGTGACATCTATAGACACCTTCGGGCGTCCCGGTGACATCGAGCTGCC	275
QY	337	AGAGGACTGATGTGTGCTGCTTCCTGATGTCCTTCTTGCTTCAATGATGECATCTT	386
Db	276	CAGGCATATAGGTGACATCCAGTGCAAATCTCTCCCTCGGCTCATATTCTGTGTG	335
QY	397	GGGATGAATGACCAAGTGCACGGGGGACATAGAAAGTGAAGGCTCACATCTGTG	456
Db	336	GGATAGATGACACACTCTTCTGCCAGGATCC--CGAGCCAAAGACAGATGCGGTA	392
QY	457	ACGGCTGATCATCTTCATCATCAAGGCGATGATGTGCTCATCCTGTGAGCTGGATT	516
Db	393	GCAGGTGAGTCTTTTTCATCTGTGAGGCGCTCGGAGTTCATCTCGTGTGCTGGAA	452
QY	517	GCCAAATGCCATCATCAGAGATTTCTATACTAAATAGTGAATGTGCCAAAAACGTAG	576
Db	453	CTTCATGGGATCTCTACGGACTTCTATCCACACTGGTCTGACAGATGAATTTTGAG	512
QY	577	CTTGGAGAAAGCTCTCACTTAGATAGAGCAACGAGCACTGATGTGTGAGAGCT	636
Db	513	ATTGGAGAGGCTTTTACTTGTGGGCAATATTTCTTCCGTGTCCTCGATAGCTGAATC	572
QY	637	CTGTTCTGCTGCGTTTTTTTGTGGCAACGAAAAGAGCACTAGCTACAGTACTGCATTCCT	686
Db	573	ATCCTCTGCTTTTCTGCTCATCCGAGAGAAATCGCTCCACTACTAGATGCTTACAA	632
QY	697	TCCGATGACAAACCAAAAAGTTATC	724
Db	633	GCCCAACTTTTCCCAACAGAGCTTC	660

```

RESULT 12
US-09-188-930-55
Sequence 55, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 55
LENGTH: 413
TYPE: DNA
ORGANISM: Human
US-09-188-930-55

```

Query Match	4.7%;	Score 94.8;	DB 3;	Length 413;
Best Local Similarity	54.6%;	Pred. No. 3.6e-15;		
Matches 189; Conservative	0;	Mismatches 157;	Indels 0;	Gaps 0

QY 295 TAAGATATCCCTGGTGGCTCTTTCTCCGAGCCATCAGGAGCAAGAGCATGATGTGTC 354

Db 21 TAGCACTGGTGTCTCCCTCTTCGCGGCTTTGCAGGCCATCGAGCCCTAAATGTGTGTC 80

QY 355 GCTTCGATGTCTCTTCTTGGCTTTCAATGAGGCATCCTTGGCATGAATCACCAG 414

Db 81 TCCCTGGTGTGGGCTTCTCGGCATGTTTGTGGCAAGATGGGCATGAAGTGCAGCGCC 140

QY 415 TGCAAGGGGAGCAATGAGGAAGGTGAAGGCTCAATTTGCTGACGCTGGAATCATCTTC 474

Db 141 TGTGGGGAGACGACAAAGTGAAGAAAGGCGGTATAGCCATGCGTGTGAAGGCATTAATTTTC 200

QY 475 ATCATCAGGGACATGGTGTGCTCATCCCTGTGAAGCTGGGTGGCAATGCCATCATCAACA 534

Db	201	ATGTGGCAGGCTTTGGCCGCTTGGTACCTTGTCTCTGGTATGGCCATCAGATTGTACA	260
Qy	535	GATTCTATATACATCATATGTGAATTGSCCCAAAACGTGAGCTTGAGAGAAGCTCTTAC	594
Db	261	GACTTTTATTAACCTTTTGAATCCCTCCAAACATTAAGTATGAGTTTGGSCCCTGCATCTT	320
Qy	595	TTAGATGACCAACGACACTGGTGTGATGTGTGAGGAGGCTCTGT	640
Db	321	ATTGGCTGGGCAAGGCTCTCCCTACTGCATCCCTGGAGAGTGCACTGT	366

```

RESULT 13
US-09-312-283C-55
; Sequence 55, Application US/09312283C
; Patent No. 657095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Marison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-55

```

Query Match	4.7%;	Score 94.8;	DB 4;	Length 413;
Best Local Similarity	54.6%;	Pred. No. 3.6e-15;		
Matches 189;	Conservative	0;	Mismatches 157;	Indels 0;
			Gaps	0;

Oy	295	TATGATATCCCTGCTGGCTCTTTCTCCGAGCTACAGGACAGGACAAGGACATATATGTC	3
Db	21	TAGAGCTCGGTGCTGCCTCTGCTCCGAGCTTCAGGAGCCATATAGTGTGTC	80
Oy	355	GCTTCCGCTATGTCTTTTGGCTTTCAATGAGCGATCCCTGGCATGAATATGACACAGG	414
Db	81	TCCCTGGTGTGGGCTTCTCGCCATGTTTTGTGGCACAGATGGGCATGAATGACACGGC	140
Oy	415	TGCACGGGAGACATATGAGAGGTGAAGGCTCACTTCTGTACGGCTGATCATCTTC	474
Db	141	TGTGGGGAGACAGACAAATGAAAGAGGCCCTATATAGCCATGGGTGAGGACATATTTTC	200
Oy	475	ATCATACGGGACATGGTGTGCTCATCCCTGTGAAGCTGGATGGCAATGCCATATAGA	534
Db	201	ATGTGGACAGTCTTCCGCTTGGTAGTGTCTCTGTATGGCATGATATGTTCACA	260
Oy	535	GATTTCTATACTCATATGATGATGTGGCCAAAACGTGAGCTTGGAGAAAGCTCTTAC	594
Db	261	GACTTTATATACCTTTGATCCCTCCAAACATTAAGTATGATTTGGCCCTGCATCTTT	320
Oy	595	TTTGAATGACACACGGGCACTGGAGCTGATGTGTGAGAGAGCTCTGT	640
Db	321	ATTGGCTGGACAGGCTGTGCTCTAGTCACTCTGGAGAGGTGACACTGT	366

RESUB: 14
US-09-404-879A-104
Sequence 104, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.


```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-104

```

```

Query Match      4.5%; Score 89.8; DB 4; Length 441;
Best Local Similarity 59.7%; Pred. No. 7.4e-14;
Matches 151; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

```

```

QY 92 GGATATGGCAACCATGCTTAGAAATCGCTGGGCTGTTCTTGTGTTGGAATGG 151
DB 178 GAACATGGCTCTCCATGGGGCTACAGTAATGGGCATCGCCTGCGCTGCGC 237
QY 152 TGGGACAGTGGCTGTCACTGATCCTCACTGAGAGAGTGTGCGCTTCATTGAAACA 211
DB 238 TGGCCGTATGCTGTGCTGGCGGCTGCCATGTGGCGCGTGAACGCTTCATGGCAGCA 297
QY 212 ACATGCTGTTTAAACCTTCTGGGAAGAGCTGTGATGAATTGGCTGAGGAGGCTA 271
DB 298 ACATGTCACTCGAGACCATCTGGAGGGCTTATGATGAATGCTGTGTGAGAGCA 357
QY 272 ACATGAGATGAGTGAATAATCTATGATTCCTGCTGGCTCTTTCTCCGAGCTACAG 331
DB 358 CCGGCAGATGAGTGAAGTGTGATGCTGCTGTGCACTGCCAGAGACCTGCAAG 417
QY 332 CAGCCAGAGGACT 344
DB 418 CGGCCCGGCCCT 430

```

```

RESULT 15
US-09-338-933-104
; Sequence 104, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 104
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-104

```

```

Query Match      4.5%; Score 89.8; DB 4; Length 441;
Best Local Similarity 59.7%; Pred. No. 7.4e-14;
Matches 151; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

```

```

QY 92 GGATATGGCAACCATGCTTAGAAATCGCTGGGCTGTTCTTGTGTTGGAATGG 151
DB 178 GAACATGGCTCTCCATGGGGCTACAGTAATGGGCATCGCCTGCGCTGCGC 237
QY 152 TGGGACAGTGGCTGTCACTGATCCTCACTGAGAGAGTGTGCGCTTCATTGAAACA 211
DB 238 TGGCCGTATGCTGTGCTGGCGGCTGCCATGTGGCGCGTGAACGCTTCATGGCAGCA 297
QY 212 ACATGCTGTTTAAACCTTCTGGGAAGAGCTGTGATGAATTGGCTGAGGAGGCTA 271
DB 298 ACATGTCACTCGAGACCATCTGGAGGGCTTATGATGAATGCTGTGTGAGAGCA 357

```

```

QY 272 ACATGAGATGCAATGCAAAAATCTATGATTCCCTGCTGCTCTTTCTCCGAGCTACAG 331
DB 358 CCGGCAGATGCAATGCAAGGTGTACGACTGCTGGGCACTGCGCAGAGACCTGCAAG 417
QY 332 CAGCCAGAGGACT 344
DB 418 CGGCCCGGCCCT 430

```

Search completed: September 2, 2004, 23:34:11
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 21:14:05 ; Search time 5283 Seconds

(without alignments)
16490.535 Million cell updates/sec

Title: US-10-063-732-119

Perfect score: 2010

Sequence: 1 ggaataactgtctctctctg.....aattgacatttttctaatt 2010

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pac.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pac.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hgco_hum.*
40: em_hgco_mus.*
41: em_hgco_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	100.0	2010	6	AX092388
2	2010	100.0	2010	6	AX376290
3	2010	100.0	2010	6	AX697259
4	2010	100.0	2010	6	AX358707
5	2008.4	99.9	137955	9	AP000884
6	2008.4	99.9	340000	9	AP001707
7	1944	96.7	267172	9	BS000177
8	1935	96.3	176580	9	BS000178
9	1929.4	96.0	1931	9	HS4250711
10	1802	89.7	1890	6	BD157589
11	1802	89.7	1890	6	BD157589
12	1802	89.7	1890	9	AC022269
13	1675.8	83.4	1700	9	BC058004
14	1550.4	77.1	1835	9	EC020866
15	1352.8	67.3	191923	2	AP001846
16	1054.8	52.5	1090	9	HS4600334
17	950	47.3	1002	6	AX497202
18	902.2	44.9	81583	2	AC131927
19	845.6	42.1	240479	2	AC096377
20	789.2	39.3	2325	10	BC003668
21	784.6	39.0	217278	2	AC113180
22	784.6	39.0	256720	10	AC110241
23	761	37.9	837	6	AX866683
24	761	37.9	837	6	BD146745
25	535.8	26.7	577	6	AX872185
26	535.8	26.7	577	6	BD152247
27	494	24.6	678	6	E31609
28	494	24.6	678	10	AF087826
29	493.6	24.6	191923	2	AP001846
30	493.2	24.5	615	6	AR414198
31	493.2	24.5	615	6	BD109751
32	365.4	18.2	429	6	AX203264
33	308.2	15.3	368	6	BD204000
34	277.6	13.8	675	9	HS4250712
35	277.6	13.8	833	6	AX497204
36	264.6	13.2	1160	6	AX697263
37	264.6	13.2	1160	9	AY358094
38	210.4	10.5	1449	10	BC058186
39	210.4	10.5	142656	2	AC133582
40	210.4	10.5	220013	10	AC122821
41	209.4	10.4	1167	9	HS4130941
42	209.4	10.4	1231	6	AX058215
43	209.4	10.4	2050	9	AK091002
44	209.4	10.4	43411	9	AC004643
45	209.4	10.4	180354	9	AC108134

ALIGNMENTS

RESULT 1
AX092388
LOCUS
DEFINITION Sequence 119 from Patent WO0116318.
ACCESSION AX092388
VERSION AX092388.1 GI:13444508
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Baton, D.L., Filvaroff, E., Gerritsen, M.E., Goddard, A., Goddard, A.,
Godowski, P.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K. and
Wood, W.I.

TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 016318-A 119 08-MAR-2001;
Genentech, Inc. (US)
FEATURES location/Qualifiers
source 1..2010
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2010; Db 6; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGAAGAACTGTTCTCTTCTGTCACAGAGAACCTGCTTCAAGCAGAAAGTACGATTC 60
1 GGAAGAACTGTTCTCTTCTGTCACAGAGAACCTGCTTCAAGCAGAAAGTACGATTC 60
61 CGAGTCACAGCTGCTTAACTCACTCCAGAGATTAAGCAACCCATCCCTTAAAGATC 120
61 CGAGTCACAGCTGCTTAACTCACTCCAGAGATTAAGCAACCCATCCCTTAAAGATC 120
121 GCTGGCGTGTCTTGT 180
121 GCTGGCGTGTCTTGT 180
181 CAGTGTGAGATGT 240
181 CAGTGTGAGATGT 240
181 CAGTGTGAGATGT 240
241 GGAATGT 300
241 GGAATGT 300
301 TCCCTGCTGCTCTTCTTCCGACCTTACAGAGACCCAGAGATGTGTGTGTGTGTGT 360
301 TCCCTGCTGCTCTTCTTCCGACCTTACAGAGACCCAGAGATGTGTGTGTGTGTGT 360
301 TCCCTGCTGCTCTTCTTCCGACCTTACAGAGACCCAGAGATGTGTGTGTGTGTGT 360
361 GTGATGTCTCTTCTGCTTCAAGATGAGCATCTTGTGCAATGAAATGCAAGATGTC 420
361 GTGATGTCTCTTCTGCTTCAAGATGAGCATCTTGTGCAATGAAATGCAAGATGTC 420
361 GTGATGTCTCTTCTGCTTCAAGATGAGCATCTTGTGCAATGAAATGCAAGATGTC 420
421 GGGGACATGAGAGGTGAAAGCTCACTTCTGCTGACGGCTGGAATCATCTTCAATC 480
421 GGGGACATGAGAGGTGAAAGCTCACTTCTGCTGACGGCTGGAATCATCTTCAATC 480
421 GGGGACATGAGAGGTGAAAGCTCACTTCTGCTGACGGCTGGAATCATCTTCAATC 480
481 ACGGGCAGT 540
481 ACGGGCAGT 540
481 ACGGGCAGT 540
541 TATACTCAATAGTGAATGTTGCCCAAAAAGTGAAGCTTGAAGAGCTCTTACTTAA 600
541 TATACTCAATAGTGAATGTTGCCCAAAAAGTGAAGCTTGAAGAGCTCTTACTTAA 600
541 TATACTCAATAGTGAATGTTGCCCAAAAAGTGAAGCTTGAAGAGCTCTTACTTAA 600
601 TGAACACAGGACCTGT 660
601 TGAACACAGGACCTGT 660
601 TGAACACAGGACCTGT 660
661 AACGAAAGAGCAGTACCTACAGTACCTGATACCTTCCATGCAACCCCAAAAAGT 720
661 AACGAAAGAGCAGTACCTACAGTACCTGATACCTTCCATGCAACCCCAAAAAGT 720
661 AACGAAAGAGCAGTACCTACAGTACCTGATACCTTCCATGCAACCCCAAAAAGT 720
721 TATACACCGGAGAGAGTCAACGAGGCTTACTCCAGAGTCAATGTGTGTGTGTGT 780
721 TATACACCGGAGAGAGTCAACGAGGCTTACTCCAGAGTCAATGTGTGTGTGTGT 780
721 TATACACCGGAGAGAGTCAACGAGGCTTACTCCAGAGTCAATGTGTGTGTGTGT 780
781 ATGTTTTTTTAACTTATTAAGCATGCAATGCAAAAATCTATTTTACTTCTTCA 840
781 ATGTTTTTTTAACTTATTAAGCATGCAATGCAAAAATCTATTTTACTTCTTCA 840
781 ATGTTTTTTTAACTTATTAAGCATGCAATGCAAAAATCTATTTTACTTCTTCA 840
841 AATGAGACCCCAAGAAAGTGAATTTACTGTTCTTAACTGCTTAACTTAACTAAG 900
841 AATGAGACCCCAAGAAAGTGAATTTACTGTTCTTAACTGCTTAACTTAACTAAG 900

Db 841 AATGAGACCCCAAGAAAGTGAATTTACTGTTCTTAACTGCTTAACTTAACTAAG 900
Qy 901 ACTGTCATCAGCTATTTATGATTTCTATAGCTATTTCAAGAGATGATATTAACC 960
Db 901 ACTGTCATCAGCTATTTATGATTTCTATAGCTATTTCAAGAGATGATATTAACC 960
Qy 961 AATGCTTATGTTCTGAAAGATATGATATTTTCTTCAAGTGTGTCAAGCATCTA 1020
Db 961 AATGCTTATGTTCTGAAAGATATGATATTTTCTTCAAGTGTGTCAAGCATCTA 1020
Qy 1021 CTCTTTTATATCTTAACTTCAAAATGACATGCTAAGAGCTGATTTTACTGCTGA 1080
Db 1021 CTCTTTTATATCTTAACTTCAAAATGACATGCTAAGAGCTGATTTTACTGCTGA 1080
Qy 1081 ATTTTCCAGCAGATAGCATTTATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 ATTTTCCAGCAGATAGCATTTATGATGATGATGATGATGATGATGATGATGATG 1140
Qy 1141 GACATGCTATATGTTTATTTATTAATGAATGCAATGCAATGCAATGCAATGCA 1200
Db 1141 GACATGCTATATGTTTATTTATTAATGAATGCAATGCAATGCAATGCAATGCA 1200
Qy 1201 ACTCACTATGCTTTTCAAGGAAATCATGATAGGCTTGAAGAGTACTATTAATG 1260
Db 1201 ACTCACTATGCTTTTCAAGGAAATCATGATAGGCTTGAAGAGTACTATTAATG 1260
Qy 1261 TTTAAAGACGCTTAGGATTAATGCTCCATTTATTAATGAAGTAAATGAAGCTT 1320
Db 1261 TTTAAAGACGCTTAGGATTAATGCTCCATTTATTAATGAAGTAAATGAAGCTT 1320
Qy 1321 TTAATGCACTTGAAGAAATGATGATGCTTCTGATATGCTGTTTTTATGCTTAAG 1380
Db 1321 TTAATGCACTTGAAGAAATGATGATGCTTCTGATATGCTGTTTTTATGCTTAAG 1380
Qy 1381 TTAGAAATCTTAACTTCTTATCTCTCTTCCAGAGCTTTTCTTCTGCTGATTTAA 1440
Db 1381 TTAGAAATCTTAACTTCTTATCTCTCTTCCAGAGCTTTTCTTCTGCTGATTTAA 1440
Qy 1441 ATTACATTTTAAACGAGATTTTGTGCAAGGCTTTGCAATCAATCTGCTTTTCC 1500
Db 1441 ATTACATTTTAAACGAGATTTTGTGCAAGGCTTTGCAATCAATCTGCTTTTCC 1500
Qy 1501 AGGGCTATCTCAAGAAAGATTAAGTGTATTAAGAAAGTATGATGTTTAAAGAA 1560
Db 1501 AGGGCTATCTCAAGAAAGATTAAGTGTATTAAGAAAGTATGATGTTTAAAGAA 1560
Qy 1561 AGTGAAGATTTTGTGTTGATTTGAGAGATGATGATGATGATGATGATGATGAT 1620
Db 1561 AGTGAAGATTTTGTGTTGATTTGAGAGATGATGATGATGATGATGATGATGAT 1620
Qy 1621 ATATGATGATATATTTTATTAAGTATTTGATGATGATGATGATGATGATGAT 1680
Db 1621 ATATGATGATATATTTTATTAAGTATTTGATGATGATGATGATGATGATGAT 1680
Qy 1681 AATTAAGAGCAGAAATATGCTGTTTCAATGCTTACCAAAAACCAACAA 1740
Db 1681 AATTAAGAGCAGAAATATGCTGTTTCAATGCTTACCAAAAACCAACAA 1740
Qy 1741 AAAAGTGTCTTTGAGAACTTCACTGCTCTTATGAGGATGCTGATGATGATGAT 1800
Db 1741 AAAAGTGTCTTTGAGAACTTCACTGCTCTTATGAGGATGCTGATGATGATGAT 1800
Qy 1801 AATTTGTTCTGAGAAATTAATTTCTTCTTGAACATTTCTGTTAGTTTAACT 1860
Db 1801 AATTTGTTCTGAGAAATTAATTTCTTCTTGAACATTTCTGTTAGTTTAACT 1860
Qy 1861 ATCTGTAATTAATCTTATTTTCTGTTTATTTCAATTTGAGAACTGATCAATTT 1920
Db 1861 ATCTGTAATTAATCTTATTTTCTGTTTATTTCAATTTGAGAACTGATCAATTT 1920
Qy 1921 GAAAGTTGTGTGAGCTGCTGCTGCTTAAATGATGATGATGATGATGATGAT 1980
Db 1921 GAAAGTTGTGTGAGCTGCTGCTGCTTAAATGATGATGATGATGATGATGAT 1980


```

Db      1621 ATATGATGATGATATATTTATATAGATTTGAGTACAGACTTTGAGTTTCATCATATA 1680
Qy      1681 AATTAAGAGCAGAAAATATGTCTTGTTTCATTGCTTACCAAAAACAAACAACA 1740
Db      1681 AATTAAGAGCAGAAAATATGTCTTGTTTCATTGCTTACCAAAAACAAACAACA 1740
Qy      1741 AAAAGGTGCTCTTGAGAACTTCACTGCTCTTATGTGGGAGCTGAGTCAAAATGTG 1800
Db      1741 AAAAGGTGCTCTTGAGAACTTCACTGCTCTTATGTGGGAGCTGAGTCAAAATGTG 1800
Qy      1801 ATTTTGTCTGTGAAAAAATTAATTCCTTCTTGTAACATTTCTGTTTATGTTTACTTAA 1860
Db      1801 ATTTTGTCTGTGAAAAAATTAATTCCTTCTTGTAACATTTCTGTTTATGTTTACTTAA 1860
Qy      1861 ATCTGTAATATGCTATTTTCTGTTTATGCAAAATTTGATGAACCTGACATCAATTT 1920
Db      1861 ATCTGTAATATGCTATTTTCTGTTTATGCAAAATTTGATGAACCTGACATCAATTT 1920
Qy      1921 GAAAGTTTGTGTCAGCTGCTGCTACCTTAATGAATGATGTTCTATTTGCTTTATACAT 1980
Db      1921 GAAAGTTTGTGTCAGCTGCTGCTACCTTAATGAATGATGTTCTATTTGCTTTATACAT 1980
Qy      1981 TATTTTATTAATTAATTTGCTTTTCTTAAT 2010
Db      1981 TATTTTATTAATTAATTTGCTTTTCTTAAT 2010

RESULT 3
AX697259      2010 bp      DNA      linear      PAT 02-APR-2003
LOCUS      AX697259
DEFINITION      Sequence 327 from Patent W00078961.
ACCESSION      AX697259
VERSION      AX697259.1 GI:29498418
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1. Ferreira, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, J.,
    Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A.,
    Godowski, P.J., Gurney, A.L., Smith, V., Tumas, D., Wood, W.I.,
    Grimaldi, C.J., Hillier, K.J., Paoni, N.F., Roy, M.A. and Matenabe, C.K.
    Secreted and transmembrane polypeptides and nucleic acids encoding
    the same
    Patent: WO 0078961-A 327 28-DEC-2000;
    Genentech Inc. (US)
FEATURES
  source
    1..2010
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 2010; DB 6; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 GAAAAAAGTCTCTCTCTCTGTCGACAGAAACCTGCTTCAACACAGAAAGTACGTC 60
Db      1 GAAAAAAGTCTCTCTCTCTGTCGACAGAAACCTGCTTCAACACAGAAAGTACGTC 60
Qy      61 CGAGAGTCAGCTGCTTAAATCTATCCAGAGATATGGAACCAACCATGCTTGAAGATC 120
Db      61 CGAGAGTCAGCTGCTTAAATCTATCCAGAGATATGGAACCAACCATGCTTGAAGATC 120
Qy      121 GCTGGGCTGTTCTTGTTGAGTGTGATGATGAGGACAGTGGCTGTCACTGTCACTGCT 180
Db      121 GCTGGGCTGTTCTTGTTGAGTGTGATGATGAGGACAGTGGCTGTCACTGTCACTGCT 180
Qy      181 CAGTGAAGAGTGTGAGCTTCAATGAAAACAATGCTGTTTGAAGAACTTGGGAA 240
Db      181 CAGTGAAGAGTGTGAGCTTCAATGAAAACAATGCTGTTTGAAGAACTTGGGAA 240

```

```

Qy      241 GGAAGTGAATGAATGCGTGAAGCAGGCTTAACATCAGATGAGTGAACAAATCTATGAT 300
Db      241 GGAAGTGAATGAATGCGTGAAGCAGGCTTAACATCAGATGAGTGAACAAATCTATGAT 300
Qy      301 TCCGCTGAGCTCTTTCTCGGAACTTACAGGACCCAGAGAGCTGATGTGTCTCTCC 360
Db      301 TCCGCTGAGCTCTTTCTCGGAACTTACAGGACCCAGAGAGCTGATGTGTGTCTCC 360
Qy      361 GGTATGCTCTTGGCTTGTATGATGAGCCATCCTTGGCATGAAATGACACAGTGCAG 420
Db      361 GGTATGCTCTTGGCTTGTATGATGAGCCATCCTTGGCATGAAATGACACAGTGCAG 420
Qy      421 GGGGCAATGAGAAAGTGAAGGCTCAATCTGCTGACGCGCTGGAATCATCTTCATC 480
Db      421 GGGGCAATGAGAAAGTGAAGGCTCAATCTGCTGACGCGCTGGAATCATCTTCATC 480
Qy      481 AGGGCAATGAGTGTCTTATCCCTGAGCTGGGTTGSCAATGCAATCATCAGAGATTG 540
Db      481 AGGGCAATGAGTGTCTTATCCCTGAGCTGGGTTGSCAATGCAATCATCAGAGATTG 540
Qy      541 TATTAATCAATGATGATGTTGCCAAAAAGTGAAGCTGGAAGAGCTCTCTATTAGA 600
Db      541 TATTAATCAATGATGATGTTGCCAAAAAGTGAAGAGCTCTCTATTAGA 600
Qy      601 TGAACCAAGGCACTGCTGCTATTTGTGAGAGAGCTCTGTTCTGCGCTTTTGTGTC 660
Db      601 TGAACCAAGGCACTGCTGCTATTTGTGAGAGAGCTCTGTTCTGCGCTTTTGTGTC 660
Qy      661 AACGAAAGAGCAGTACAGATACATCTCATCTTCCATGCGCAACCCAAAAAGT 720
Db      661 AACGAAAGAGCAGTACAGATACATCTCATCTTCCATGCGCAACCCAAAAAGT 720
Qy      721 TATCAACCCGAAAGAGTCAACCGAGGCTCTATCTCCAGAAAGTCAGTATGTGTGTGT 780
Db      721 TATCAACCCGAAAGAGTCAACCGAGGCTCTATCTCCAGAAAGTCAGTATGTGTGTGT 780
Qy      781 ATGTTTTTTTAACTTACTATTAAGCCATGCAATGACAAATATCTATTACTTTCTCA 840
Db      781 ATGTTTTTTTAACTTACTATTAAGCCATGCAATGACAAATATCTATTACTTTCTCA 840
Qy      841 AATGAGCCCCAAGAAACCTTGAATTTACTGTTCTTAATCTGCTTAATTTACAGA 900
Db      841 AATGAGCCCCAAGAAACCTTGAATTTACTGTTCTTAATCTGCTTAATTTACAGA 900
Qy      901 ACTGTGATCAGCTATTTATATCTATTAAGCTATTTCAAGAGATGAGATTTAAACC 960
Db      901 ACTGTGATCAGCTATTTATATCTATTAAGCTATTTCAAGAGATGAGATTTAAACC 960
Qy      961 AATGCTTGAATGTTCTGAAAGATAGTAAATTTGTTTCTAAGTGTCAAGCATCTA 1020
Db      961 AATGCTTGAATGTTCTGAAAGATAGTAAATTTGTTTCTAAGTGTCAAGCATCTA 1020
Qy      1021 CTTCTTTATCTATTTACTTCAAAATGACATGCTGAAGACCTGATTTTACTACTGTA 1080
Db      1021 CTTCTTTATCTATTTACTTCAAAATGACATGCTGAAGACCTGATTTTACTACTGTA 1080
Qy      1081 ATTTCTCAGCAGATAGCATATGATGATAGTGAACATTTATATCTCAGATAGA 1140
Db      1081 ATTTCTCAGCAGATAGCATATGATGATAGTGAACATTTATATCTCAGATAGA 1140
Qy      1141 GACATGCTTAATGTTTAAATGAAATGCAATGCAATTTATGCTGAATTAAGA 1200
Db      1141 GACATGCTTAATGTTTAAATGAAATGCAATGCAATTTATGCTGAATTAAGA 1200
Qy      1201 ACTCAATATGCTTTTCAAGGAAATCATGATGAGGTTGAAGAGTCTATTAATG 1260
Db      1201 ACTCAATATGCTTTTCAAGGAAATCATGATGAGGTTGAAGAGTCTATTAATG 1260
Qy      1261 TTTAAAAACAGCTTGAAGATTAATGCTCTCATTTATATGAAGATTTAAAGAGCTT 1320
Db      1261 TTTAAAAACAGCTTGAAGATTAATGCTCTCATTTATATGAAGATTTAAAGAGCTT 1320

```

Qy	132	TAAATCGACATGAAAGAAATGAAATGGCTTCTGATATGCTGTTTTTACCTAGAG	1380
Db	1321	TAAATCGACATGAAAGAAATGAAATGGCTTCTGATATGCTGTTTTTACCTAGAG	1380
Qy	1381	TTGAAGATCCTAAGCTTCTTATCTCTCTCCAGAGGCTTTTTTTCTGTGATTTA	1440
Ds	1381	TTGAAGATCCTAAGCTTCTTATCTCTCTCCAGAGGCTTTTTTTCTGTGATTTA	1440
Qy	1441	ATTAAATTTTTTAAACGAGATATTTGTGCAAGGGCTTGGATCAATCGCTTTCC	1500
Ds	1441	ATTAAATTTTTTAAACGAGATATTTGTGCAAGGGCTTGGATCAATCGCTTTCC	1500
Qy	1501	AGGGCTATATCTCAGAGAAAGATATAAGTGTATCTAAGAAAAAGATGTTTTAGAA	1560
Ds	1501	AGGGCTATATCTCAGAGAAAGATATAAGTGTATCTAAGAAAAAGATGTTTTAGAA	1560
Qy	1561	AGTGAATAATTTTTGTGTTTTGTATTTGTAAGAAAGATATGCAATTTTGACAGAAATCAT	1620
Ds	1561	AGTGAATAATTTTTGTGTTTTGTATTTGTAAGAAAGATATGCAATTTTGACAGAAATCAT	1620
Qy	1621	ATATGATGATATATTTTAAATAGATATTTGATGATACAGCTTTGAGGTTTCATCAATTA	1680
Ds	1621	ATATGATGATATTTTAAATAGATATTTGATGATACAGCTTTGAGGTTTCATCAATTA	1680
Qy	1681	AATAAAGAGCGAGAAAATATGTCTGTGTTTCATTGCTTACCAAAAAACAACA	1740
Ds	1681	AATAAAGAGCGAGAAAATATGTCTGTGTTTCATTGCTTACCAAAAAACAACA	1740
Qy	1741	AAAAAGTGTCCCTTGAAGACTTCAACGCTGCTCATGTGGGTACCTGAGTCAAAATGTC	1800
Ds	1741	AAAAAGTGTCCCTTGAAGACTTCAACGCTGCTCATGTGGGTACCTGAGTCAAAATGTC	1800
Qy	1801	ATTTTGTGTCTGTGAATAATAATTTCTTGTGTACATTTCTGTTAATTTCTAA	1860
Ds	1801	ATTTTGTGTCTGTGAATAATAATTTCTTGTGTACATTTCTGTTAATTTCTAA	1860
Qy	1861	ATCTGAATAATCGATATTTTCTGTTATTTGCAAAATGATGATGAACTGACATTCATTT	1920
Ds	1861	ATCTGAATAATCGATATTTTCTGTTATTTGCAAAATGATGATGAACTGACATTCATTT	1920
Qy	1921	GAAAGTTGTGTGACAGTCTGTCTAGAGCTTAATGAATGTTTCTATTTGCTTTATACAT	1980
Ds	1921	GAAAGTTGTGTGACAGTCTGTCTAGAGCTTAATGAATGTTTCTATTTGCTTTATACAT	1980
Qy	1981	TATATTAATTAATGACATTTTCTAAT	2010
Ds	1981	TATATTAATTAATGACATTTTCTAAT	2010

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	REFERENCE	AUTHORS	TITLE
AY358707	AY358707	Homo sapiens clone DNA3735 C1DN8 (NM0779)	2010 bp	mrna	linear	ERI 03-OCT-2003			
AY358707	AY358707.1	GI:37182535							
	FLI CDNA.								
	Homo sapiens								
	Homo sapiens (human)								
	Bakayoceta; Metazoa; Chordata; Cranitca; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.								
	1 (bases 1 to 2010)								
	Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Denel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowicz,L., Jin,T., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V., Sincison,J., Vagts,A., Vandlen,R., Wetanabe,C., Wiand,D., Woods,K., Xie,M.H., Yanaura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,M.I. and Godowski,P.								
	The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment								

JOURNAL Genome Res. 13 (10), 2265-2270 (2003)
 PUBMED 12975309
 REFERENCE 2 (bases 1 to 2010)
 AUTHORS Clark,H.F.
 TITLE Direct Submision
 JOURNAL Submitted (01-AUG-2003) Department of Bioinformatics, Genentech
 Inc., 1 DNA way, South San Francisco, CA 94080, USA
 FEATURES Location/Qualifiers

gene	1. .2010
CDS	/locus_tag="UNQ779" 97. .774

```

/terminal="MATHALEINFLIGVGWGVAVLTVAPORVASFIENNM
/ctnslcmmnccroantimocktydslls.plolaasmsl.pama.ii.com
kctrcgtednekvkhililngiifitkvyulipswanaki.rpfnslynaokce
lgealvlgmwtlavlvgal.fccvpcnckssysrslpshrtoks.yhnrkkspsv
ysrcqyv"

```

Query Match	100.0%	Score 2010	DB 9	Length 2010
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2010	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
QY	1	GGAAAACTGTTCTCTCTGTGTGGCAAGAACCTCTGCTTAAAGAGAGATGACATTC	60	
Db	1	GGAAAAACGTGTTCTCTCTGTGGCAGAGAACCTCTGCTTCAACAGAGATGACATTC	60	
QY	61	CGAGTCACAGCTGGCTAAATCATCCACAGGATTAATGGCAACCATGCTTGAATTC	120	
Db	61	CGAGTCACAGCTGGCTAAATCATCCACAGGATTAATGGCAACCATGCTTGAATTC	120	
QY	121	GCTGGGCTGTTCTTGATGATGTTGAATGTGGTGGCAAGTGGCTGCACTGTACGCTT	180	
Db	121	GCTGGGCTGTTCTTGATGATGTTGAATGTGGTGGCAAGTGGCTGCACTGTACGCTT	180	
QY	181	CAGTGGAGATGTGGGCTTCAATTGAAAACAATGTGTGTTTGAATTCCTGGGAA	240	
Db	181	CAGTGGAGATGTGGGCTTCAATTGAAAACAATGTGTGTTTGAATTCCTGGGAA	240	
QY	241	GGACTGTGATGAATTGCGTGAAGGACGGCTTAACATCAGAGTSCAGTGCATAATCAT	300	
Db	241	GGACTGTGATGAATTGCGTGAAGGACGGCTTAACATCAGAGTSCAGTGCATAATCAT	300	
QY	301	TGCCGTGCGCTCTTCTTCCGGAACATACAGGACGCCAGAGACATGATGTCGCTTCC	360	
Db	301	TGCCGTGCGCTCTTCTTCCGGAACATACAGGACGCCAGAGACATGATGTCGCTTCC	360	
QY	361	GTGATGTCTTCTTGCGCTTTCATGATGGCCATCTTGGCATGAATGACAGGTGCACG	420	
Db	361	GTGATGTCTTCTTGCGCTTTCATGATGGCCATCTTGGCATGAATGACAGGTGCACG	420	
QY	421	GGGACATGAGAGGTGAAGGCTCAATCTGTCGACGGCTGGAATCATCTTCAATC	480	
Db	421	GGGACATGAGAGGTGAAGGCTCAATCTGTCGACGGCTGGAATCATCTTCAATC	480	
QY	481	ACGGGCATGTGTGCTCATCCGTGAGCTGGGTTGCCAATGCCATCATGAGATTC	540	
Db	481	ACGGGCATGTGTGCTCATCCGTGAGCTGGGTTGCCAATGCCATCATGAGATTC	540	
QY	541	TATACTCAATATGTAATGTTGCCAAAAACGTGAGCTTGGAGAGCTCTCTACTTAGA	600	
Db	541	TATACTCAATATGTAATGTTGCCAAAAACGTGAGCTTGGAGAGCTCTCTACTTAGA	600	

```

QY 601 TGGACCAAGGCACTGGTGTGATGTTGGAGAGCTGTGTTGCTGGGCTTTTGTGTGC 660
Db 601 TGGACCAAGGCACTGGTGTGATGTTGGAGAGCTGTGTTGCTGGGCTTTTGTGTGC 660
QY 661 AAGGAAAGAGAGTACAGATACATGATACCTTCCATCGACCAACCCAAAAGT 720
Db 661 AAGGAAAGAGAGTACAGATACATGATACCTTCCATCGACCAACCCAAAAGT 720
QY 721 TATCAACCGGAAAGATGACCGAGCGTCTACTCCAGAACTCAAGTATGTGTGTGT 780
Db 721 TATCAACCGGAAAGATGACCGAGCGTCTACTCCAGAACTCAAGTATGTGTGTGT 780
QY 781 ATGTTTTTAACTTACTATTAAGCCATGCAAAATGACAAAATCTATTTACTTTCTCA 840
Db 781 ATGTTTTTAACTTACTATTAAGCCATGCAAAATGACAAAATCTATTTACTTTCTCA 840
QY 841 AAATGACCCCAAGAACTTTGATTTACTGTTCTTAAGTCTTAATCTTAATCAAGA 900
Db 841 AAATGACCCCAAGAACTTTGATTTACTGTTCTTAAGTCTTAATCTTAATCAAGA 900
QY 901 AAGTGTGATAGCTATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGAT 960
Db 901 AAGTGTGATAGCTATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGAT 960
QY 961 AATGCTTTGATTTGTTCTAGAAAGTATGATTTGTTTCTAAGGTGCTTCAAGCATCTA 1020
Db 961 AATGCTTTGATTTGTTCTAGAAAGTATGATTTGTTTCTAAGGTGCTTCAAGCATCTA 1020
QY 1021 CTCTTTTATCATTTTACTTCAAAATGACATGCTTAAGCTGATTTTATTTACTGCTA 1080
Db 1021 CTCTTTTATCATTTTACTTCAAAATGACATGCTTAAGCTGATTTTATTTACTGCTA 1080
QY 1081 ATTCTTCACGACATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGAT 1140
Db 1081 ATTCTTCACGACATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGAT 1140
QY 1141 GACATGCTTTATGATTTTATTTAAATGAAATGCGATGCTTCACTGATTAATGTA 1200
Db 1141 GACATGCTTTATGATTTTATTTAAATGAAATGCGATGCTTCACTGATTAATGTA 1200
QY 1201 ACTCACTATGCTTTTCCAGGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 ACTCACTATGCTTTTCCAGGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 TTTAAAAAGCTTTAGGATTAATGCTCCATTTATTAAGATTAATGAAGGCTT 1320
Db 1261 TTTAAAAAGCTTTAGGATTAATGCTCCATTTATTAAGATTAATGAAGGCTT 1320
QY 1321 TTAATGCACTTTAAAGAAATGATGATGCTTCTGAATGCTGTTTATGCTTATGCTTATGCT 1380
Db 1321 TTAATGCACTTTAAAGAAATGATGATGCTTCTGAATGCTGTTTATGCTTATGCTTATGCT 1380
QY 1381 TTAGAAATCCTAATCTTTATCTCTCTCTCCAGAGGCTTTTCTTGTGTATTTA 1440
Db 1381 TTAGAAATCCTAATCTTTATCTCTCTCTCCAGAGGCTTTTCTTGTGTATTTA 1440
QY 1441 ATTAAACATTTTAAAGCAGATATTTGTGCAAGGGGCTTTGATTTCAAACTGCTTTCC 1500
Db 1441 ATTAAACATTTTAAAGCAGATATTTGTGCAAGGGGCTTTGATTTCAAACTGCTTTCC 1500
QY 1501 AGGGCTTACTCAGAAAGAAATGATGATGCTTCTAAGAAAGTGAATGATGATGATGATGAT 1560
Db 1501 AGGGCTTACTCAGAAAGAAATGATGATGCTTCTAAGAAAGTGAATGATGATGATGATGAT 1560
QY 1561 AGTGAATAATTTTGTGTTTGTATTTGATTTGAAGAAATGATGATTTTGAAGAAATGAT 1620
Db 1561 AGTGAATAATTTTGTGTTTGTATTTGATTTGAAGAAATGATGATTTTGAAGAAATGAT 1620
QY 1621 ATATGATGATATATTTTAAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 ATATGATGATATATTTTAAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1680
QY 1681 AATTAAGAGCAGAAATAATGATGCTTGTGTTTCACTTCAAAATAACACACACA 1740

```

```

Db 1681 AATTAAGAGCAGAAATAATGATGCTTGTGTTTCACTTCAAAATAACACACACA 1740
QY 1741 AAAAGTGTCTTTGAGAACTTCACTGCTCTATGTGGTACCTGATCAAAATTTGTC 1800
Db 1741 AAAAGTGTCTTTGAGAACTTCACTGCTCTATGTGGTACCTGATCAAAATTTGTC 1800
QY 1801 ATTTTGTCTGTAATAATTTCTTTGATCAATTTGATGAAATGATGAAATGATGAAAT 1860
Db 1801 ATTTTGTCTGTAATAATTTCTTTGATCAATTTGATGAAATGATGAAATGATGAAAT 1860
QY 1861 ATCTGTAATACTGATTTTCTGTTTATTCGAAATTTGATGAAATGATGAAATGATGAAAT 1920
Db 1861 ATCTGTAATACTGATTTTCTGTTTATTCGAAATTTGATGAAATGATGAAATGATGAAAT 1920
QY 1921 GAAAGTGTGTCGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 GAAAGTGTGTCGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1980
QY 1981 TATATTAAATTAATGATTTTCTAAT 2010
Db 1981 TATATTAAATTAATGATTTTCTAAT 2010

RESULT 5
AP000884/c 137955 bp DNA linear PRI 16-DEC-1999
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, clone: B463J19,
DEFINITION SOD-AML region, complete sequence.
ACCESSION AP000884
VERSION AP000884.1 GI:6580113
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 137955)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Hom sapiens 137,955 genomic DNA of 21q22.1
Published Only in Database (1999)
2 (bases 1 to 137955)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (15-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hdp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)

FEATURES
Location/Qualifiers
source 1..137955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="B463J19"

ORIGIN
Query Match 99.9%; Score 2008.4; DB 9; Length 137955;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2009; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAATAAGCTGTTCTCTTCTGTCGACACAGAAACCTGTTCAAGCAGAGTACCACTTC 60
Db 58357 GGAATAAGCTGTTCTCTTCTGTCGACACAGAAACCTGTTCAAGCAGAGTACCACTTC 58298
QY 61 CGAGTCCAGCTGCTTAATCACTCCAGAGATTAATGGCAACCATGCTTGAATC 120
Db 58297 CGAGTCCAGCTGCTTAATCACTCCAGAGATTAATGGCAACCATGCTTGAATC 58238

```


QY 121 GCTGGGCTGTTCTTGGTGTGTGGAATGTTGGGACAGTGGCTGTCATGCTGCT 180
 Db 58237 GCTGGGCTGTTCTTGGTGTGTGGAATGTTGGGACAGTGGCTGTCATGCTGCT 58178
 QY 181 CAGTGGAGAGTGTGGCTTCATTGAAAACAACGTGGTTTTTGA AAACTTCTGGAA 240
 Db 58177 CAGTGGAGAGTGTGGCTTCATTGAAAACAACGTGGTTTTTGA AAACTTCTGGAA 58118
 QY 241 GGACTGGATGATGATGGGTGAGGAGGCTAACATGAGATGAGTGAATATATATAT 300
 Db 58117 GGACTGGATGATGATGGGTGAGGAGGCTAACATGAGATGAGTGAATATATATAT 58038
 QY 301 TCCCTGCTGCTCTTCTCCGAGCTTACAGGAGCCAGAGAGCTGATGCTGCTTCC 360
 Db 58057 TCCCTGCTGCTCTTCTCCGAGCTTACAGGAGCCAGAGAGCTGATGCTGCTTCC 57938
 QY 361 GTGATGCTCTCTTGGCTTCAATGAGGCACTCTTGGCAATGAAATGCAACAGTGGCAG 420
 Db 57997 GTGATGCTCTCTTGGCTTCAATGAGGCACTCTTGGCAATGAAATGCAACAGTGGCAG 57938
 QY 421 GGGGACATGAGAGAGTGAAGGCTCACATTCTGCTGAGCGCTGGAATTCATCTTCATC 480
 Db 57937 GGGGACATGAGAGAGTGAAGGCTCACATTCTGCTGAGCGCTGGAATTCATCTTCATC 57878
 QY 481 ACGGGCATGCTGCTGCTCATCCCTGTGAGCTGGTGGCCATGCTCATCAGAGATTC 540
 Db 57877 ACGGGCATGCTGCTGCTCATCCCTGTGAGCTGGTGGCCATGCTCATCAGAGATTC 57818
 QY 541 TATTAATCAATAGTGAATGTTGCCCAAAAAGTGAAGCTTGAAGAGCTCTTACTTAA 600
 Db 57817 TATTAATCAATAGTGAATGTTGCCCAAAAAGTGAAGAGCTCTTACTTAA 57758
 QY 601 TGAACACGCGCATGCTGCTGATGTTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 57757 TGAACACGCGCATGCTGCTGATGTTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCT 57698
 QY 661 AAGGAAAAGAGCACTAGCTCAGATCTGATACCTTCCCATGGCAACCCCAAAAAT 720
 Db 57697 AAGGAAAAGAGCACTAGCTCAGATCTGATACCTTCCCATGGCAACCCCAAAAAT 57638
 QY 721 TATCACACCGGAAAAGAGTCAACCGAGCTTACTCCAGAGTCAATGATGATGATG 780
 Db 57637 TATCACACCGGAAAAGAGTCAACCGAGCTTACTCCAGAGTCAATGATGATGATG 57578
 QY 781 ATGTTTTTTTAACTTTACTATTAAGCAATGCAATGCAAAAATCTATATTTACTTTCT 840
 Db 57577 ATGTTTTTTTAACTTTACTATTAAGCAATGCAATGCAAAAATCTATATTTACTTTCT 57518
 QY 841 AAATGGAACCCCAAGAACTTTGATTTACTGTTCTTAACTGCTTAACTTAACTTAA 900
 Db 57517 AAATGGAACCCCAAGAACTTTGATTTACTGTTCTTAACTGCTTAACTTAACTTAA 57458
 QY 901 ACTGTGATCAGCTATTATGATCTATTAAGCTATTGAGAGATGAGATATTTAAACC 960
 Db 57457 ACTGTGATCAGCTATTATGATCTATTAAGCTATTGAGAGATGAGATATTTAAACC 57398
 QY 961 AATGCTTGAATGTTCTAGAAAGATATGTAATTTGTTTCTAAGTGGTCAAGCATCTA 1020
 Db 57397 AATGCTTGAATGTTCTAGAAAGATATGTAATTTGTTTCTAAGTGGTCAAGCATCTA 57338
 QY 1021 CTCTTTTATCATTTACTTCAAAAATGATGCTGTAAGAGCTGATTTATTTACTGTA 1080
 Db 57337 CTCTTTTATCATTTACTTCAAAAATGATGCTGTAAGAGCTGATTTATTTACTGTA 57278
 QY 1081 ATTTCGACAGACATACATTAATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 57277 ATTTCGACAGACATACATTAATGATGATGATGATGATGATGATGATGATGATGAT 57218
 QY 1141 GACATGCTTATGATGTTTATTTAAATGAAATGCAAGTCCATTAAGTAAATATAGA 1200
 Db 57217 GACATGCTTATGATGTTTATTTAAATGAAATGCAAGTCCATTAAGTAAATATAGA 57158
 QY 1201 ACTCAACTATTGCTTTTCAAGGAATATATGATAGGCTTGAAGAGCTTACTATTAATTG 1260

Db 57157 ACTCAACTATTGCTTTTCAAGGAATATATGATAGGCTTGAAGAGCTTACTATTAATTG 57038
 QY 1261 TTTTAAAAACAGCTTAAAGGATTAATGCTTCAATTTAATGAAGATTAATAAGAGCTT 1320
 Db 57097 TTTTAAAAACAGCTTAAAGGATTAATGCTTCAATTTAATGAAGATTAATAAGAGCTT 57038
 QY 1321 TAATCAGATTTGTAAGAAATGAATGAGCTTCTGATATGAGCTTTTAAAGCTGAG 1380
 Db 57037 TAATCAGATTTGTAAGAAATGAATGAGCTTCTGATATGAGCTTTTAAAGCTGAG 56978
 QY 1381 TTGAATATCTTAATCTTTATCTCTCTTCTCCAGAGGCTTTTTTTTCTGTGATTA 1440
 Db 56977 TTGAATATCTTAATCTTTATCTCTCTTCTCCAGAGGCTTTTTTTTCTGTGATTA 56918
 QY 1441 ATTAACTTTTAAACGAGATATTTGTAAGAGGCTTTGATTAAGAGCTTTTCC 1500
 Db 56917 ATTAACTTTTAAACGAGATATTTGTAAGAGGCTTTGATTAAGAGCTTTTCC 56858
 QY 1501 AGGCTTACTCAGAAAGAAATGAATGATGATCTTAAGAAAAGTATGATTTAGAA 1560
 Db 56857 AGGCTTACTCAGAAAGAAATGAATGATGATCTTAAGAAAAGTATGATTTAGAA 56798
 QY 1561 AGTGAATATTTTGTGTTTGTATTTGTAAGAAAGATGATGATTTTGAACAAGATCAT 1620
 Db 56797 AGTGAATATTTTGTGTTTGTATTTGTAAGAAAGATGATGATTTTGAACAAGATCAT 56738
 QY 1621 ATATGATGATATATTTTAAATAGATTTGATGATGATGATGATGATGATGATGAT 1680
 Db 56737 ATATGATGATATATTTTAAATAGATTTGATGATGATGATGATGATGATGATGAT 56678
 QY 1681 AATTAAGACAGAAAAATGCTGTTTCTTCAATTTCTTCAACAAAAACAACA 1740
 Db 56677 AATTAAGACAGAAAAATGCTGTTTCTTCAATTTCTTCAACAAAAACAACA 56618
 QY 1741 AAAAGTTCCTTTGAGAACTCACCTGCTCTATGAGGTAACCTGAGTCAAAATGTC 1800
 Db 56617 AAAAGTTCCTTTGAGAACTCACCTGCTCTATGAGGTAACCTGAGTCAAAATGTC 56558
 QY 1801 ATTTTGTCTGAAAAATTAATTTCTTCTTGAACATTTCTGTTTATGTTTACTAA 1860
 Db 56557 ATTTTGTCTGAAAAATTAATTTCTTCTTGAACATTTCTGTTTATGTTTACTAA 56498
 QY 1861 ATCTGTAATATGATTTTCTGTTTATTTCCAAATTTGATGAACAGTCAATCAATTT 1920
 Db 56497 ATCTGTAATATGATTTTCTGTTTATTTCCAAATTTGATGAACAGTCAATCAATTT 56438
 QY 1921 GAAAGTTTGTGAGAGCTGCTTACTTAAATGATGATGTTCTAATTTGCTTAAAT 1980
 Db 56437 GAAAGTTTGTGAGAGCTGCTTACTTAAATGATGATGTTCTAATTTGCTTAAAT 56378
 QY 1981 TATATTAATTAATTTGATTTTCTTAATTTCTTAATTTCTTAATTTCTTAATTT 2040
 Db 56377 TATATTAATTAATTTGATTTTCTTAATTTCTTAATTTCTTAATTTCTTAATTT 56348

RESULT 6
 AB001707/c 340000 bp DNA linear PRI 21-MAY-2003
 LOCUS Homo sapiens genomic DNA, chromosome 21q, section 51/105.
 DEFINITION AB001707 ALI63252 BA000005
 VERSION AB001707.1 GI:7768786
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Hattori, M., Fujiyama, A., Taylor, T. D., Marand, H., Yada, T.,
 Park, H. S., Toyoda, A., Ishii, K., Totsuki, Y., Choi, D. K., Sodea, E.,
 Okhi, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K.,
 Polley, A., Menzel, U., Delabre, J., Kumpf, K., Lehmann, R.,
 Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A.,

[illegible][illegible]

[illegible]

Db 294 GGAAGTGGATGAATGGCTGAGGAGGCTAACATCAGGATGCAAGTCAAAATCTATGAT 353
QY 301 TCCCTGCTGGCTCTTTCTCCGACCTACAGCAGCAGAGGACTGATGTGCTGCTTCC 360
Db 354 TCCCTGCTGGCTCTTTCTCCGACCTACAGCAGCAGAGGACTGATGTGCTGCTTCC 413
QY 361 GTGATGCTCTTCTGGCTTTTCATGATGCCATCTTGGCATGAAATGCCAGGTGCACG 420
Db 414 GTGATGCCCTTCTGGCTTTTCATGATGCCATCTTGGCATGAAATGCCAGGTGCACG 473
QY 421 GGGGCAATGAGAGGCTGAGGCTCACATCTGCTGAGCGGTGCAATCATCTTCATCATC 480
Db 474 GGGGCAATGAGAGGCTGAGGCTCACATCTGCTGAGCGGTGCAATCATCTTCATCATC 533
QY 481 ACGGGCAGTGGTGGTCTCATCTCCCTGAGCTGGGTGCGCAATGCCATCATCAGAGATTC 540
Db 534 ACGGGCAGTGGTGGTCTCATCTCCCTGAGCTGGGTGCGCAATGCCATCATCAGAGATTC 593
QY 541 TATACTCAATGATGAATGTTGCCCAAAAGCTGAGCTTGGAGAGCTCTCTACTTAGGA 600
Db 594 TATACTCAATGATGAATGTTGCCCAAAAGCTGAGCTTGGAGAGCTCTCTACTTAGGA 653
QY 601 TGGACACAGGCACTGGTGTGATGTTGGAGGAGCTCTGTTCTGCTGCGTCTTTTGTGTC 660
Db 654 TGGACACAGGCACTGGTGTGATGTTGGAGGAGCTCTGTTCTGCTGCGTCTTTTGTGTC 713
QY 661 AAGCAAAAGAGCAGTGTGATCAGATCTCGATACCTTCCCATCGCACAAACCAAAAGT 720
Db 714 AAGCAAAAGAGCAGTGTGATCAGATCTCGATACCTTCCCATCGCACAAACCAAAAGT 773
QY 721 TATCACACGGAAAGAGTCCAGGAGCTCTACTCCAGAGCTAGTGTGATGTGTTGT 780
Db 774 TATCACACGGAAAGAGTCA-----CGTCTACTCCAGAGTCAATGTGATGTGTTGT 828
QY 781 ATGTTTTTTTAACTTTACTATAAGCCATCAAAATGACAAAAATCTATTTACTTTCTCA 840
Db 829 ATGTTTTTTTAACTTTACTATAAGCCATCAAAATGACAAAAATCTATTTACTTTCTCA 888
QY 841 AATGACACCCCAAGAACTTTGATTTACTGTTCTTAACTGCTTAACTTTAACTTACAGGA 900
Db 889 AATGACACCCCAAGAACTTTGATTTACTGTTCTTAACTGCTTAACTTTAACTTACAGGA 948
QY 901 ACTGTGATCAGCTATTTATGATTTCTATAAGCTATTTTACAGCAGATGAGATTTAAACCC 960
Db 949 ACTGTGATCAGCTATTTATGATTTCTATAAGCTATTTTACAGCAGATGAGATTTAAATCC 1008
QY 961 AATGTTTTGATTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAGCATCTA 1020
Db 1009 AATGTTTTGATTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAGCATCTA 1068
QY 1021 CTCTTTTATCATTTTACTTCAAAATGACATTCCTAAAGCTGCATTTTCTACTACTGTA 1080
Db 1069 CTCTTTTATCATTTTACTTCAAAATGACATTCCTAAAGCTGCATTTTCTACTACTGTA 1128
QY 1081 ATTTCTCCACGATAGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1129 ATTTCTCCACGATAGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1188
QY 1141 GACATGCTTATATGTTTTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1200
Db 1189 GACATGCTTATATGTTTTTATTAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1248
QY 1201 ACTCAACTATGCTTTTTCAGGAAATCATGATAGGTTGAGAGGTTTACTATTAATTG 1260
Db 1249 ACTCAACTATGCTTTTTCAGGAAATCATGATAGGTTGAGAGGTTTACTATTAATTG 1308
QY 1261 TTTAAAAACAGCTTAGGATTAATGCTCTCCATTTTATATGAGATTTAAATGAGGCTT 1320
Db 1309 TTTAAAAACAGCTTAGGATTAATGCTCTCCATTTTATATGAGATTTAAATGAGGCTT 1368
QY 1321 TATACAGCTTGTAAAGGAAATGAAATGCTTCTGATATGCTGTTTTTTAGGCTTAGGAG 1380

Db 1369 TAATCAGCAATTTGTAAGAAATTTGAATGGCTTTCTGATATGCTGTTTTTTAGCCTAGGAG 1428
QY 1381 TTAGAAATCTTAACATCTTTTATCTCTCTCCAGAGGC-TTTTTTTTCTTGATTA 1439
Db 1429 TTAGAAATCTTAACATCTTTTATCTCTCTCCAGAGGC-TTTTTTTTCTTGATTA 1488
QY 1440 AATTAACTTTTAAACGCGAGATATTTTGTCAAGGGCTTTGCAATTCAACTGCTTTTC 1499
Db 1489 AATTAACTTTTAAACGCGAGATATTTTGTCAAGGGCTTTGCAATTCAACTGCTTTTC 1548
QY 1500 CAGGCTATCTACAGAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGA 1559
Db 1549 CAGGCTATCTACAGAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGA 1608
QY 1560 AAGTGAATAATTTTGTGTTTGTGATTTGAAGAAAGATGATGATTTGAAGAATAATCA 1619
Db 1609 AAGTGAATAATTTTGTGTTTGTGATTTGAAGAAAGATGATGATTTGAAGAATAATCA 1668
QY 1620 TATATGATGATATATTTTAAATAGTATTTAGTACAGACTTTGAGCTTTCATCAATAT 1679
Db 1669 TATATGATGATATATTTTAAATAGTATTTAGTACAGACTTTGAGCTTTCATCAATAT 1728
QY 1680 AATAAAGAGCAGAAAAATATGCTTGGTTTTTCATTTGCTTACCAAAAAACAACA 1739
Db 1729 AATAAAGAGCAGAAAAATATGCTTGGTTTTTCATTTGCTTACCAAAAAACAACA 1788
QY 1740 AAAAAAGTGTCTTGTGAGAACTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1799
Db 1789 AAAAAAGTGTCTTGTGAGAACTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1848
QY 1800 CATTTTTTCTCTGTGAAAAATAAATTTCTTTTGTACCAT 1841
Db 1849 CATTTTTTCTCTGTGAAAAATAAATTTCTTTCTTTGACCAT 1890

RESULT 11
BD157589
LOCUS 1890 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157589
VERSION BD157589.1 GI:27863347
KEYWORDS JP 2002191363-A/12432.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1890)
AUTHORS Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 12432 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/12432
PD 09-JUL-2002 JP 2002280990
PF 28-JUL-2000 JP 2002280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers (150)..(848).
FT CDS Location/Qualifiers
1..1890
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Sugawara M., Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Ninomiya K. and Iwayanagi T.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 1890)
Isogai T. and Otsuki T.
Direct Submission
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- and 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
Location/Qualifiers
1..1890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1000956"
/tissue_type="Mammary gland"
/clone_lib="MAMMA1"
/note="Cloning vector: pME18SFL3"
misc_feature
795..796
/note="5 bases segment is present in AJ250711, AP001707.1 and AP000884.1."
ORIGIN
Query Match 89.7%; Score 1802; DB 9; Length 1890;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1831; Conservative 0; Mismatches 5; Indels 6; Gaps 2;
QY 1 GGAAGAACTGTTCTTCTGTGGCAGAGAACCTGCTTCAAGCAGAGAGTAGCAGTTTC 60
DB 54 GGAAGAACTGTTCTTCTGTGGCAGAGAACCTGCTTCAAGCAGAGAGTAGCAGTTTC 113
QY 61 CGGAGTCAGCTGCTAAATCATCCAGAGGATATGCAACCCATGCTTAGAAATC 120
DB 114 CGGAGTCAGCTGCTAAATCATCCAGAGGATATGCAACCCATGCTTAGAAATC 173
QY 121 GCTGGGCTGTTTCTTGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
DB 174 GCTGGGCTGTTTCTTGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 233
QY 181 CAGTGGAGAGTGTGGGCTTCATTGAAAACAAATCGTGGTGGTGGTGGTGGTGGTGGT 240
DB 234 CAGTGGAGAGTGTGGGCTTCATTGAAAACAAATCGTGGTGGTGGTGGTGGTGGTGGT 293
QY 241 GGACTGTGGATGAATGTGGTGGGAGGCTAAATCAGGATGCAATGCAATGCAATGCAAT 300
DB 294 GGACTGTGGATGAATGTGGTGGGAGGCTAAATCAGGATGCAATGCAATGCAATGCAAT 353
QY 301 TCCCTGTGGCTGTTTCTCCGACCTACAGGAGCAGAGAGCTGATGTGTGTGTGTGTTC 360
DB 354 TCCCTGTGGCTGTTTCTCCGACCTACAGGAGCAGAGAGCTGATGTGTGTGTGTGTTC 413
QY 361 GTGATGTCTTCTGGCTTCATGATGGCCATCTTGGCATGAATCACCAGGTGCAAG 420
DB 414 GTGATGTCTTCTGGCTTCATGATGGCCATCTTGGCATGAATGCAATGCAATGCAAT 473
QY 421 GGGGCAATGAGAAGGTGAAGGCTCACAATCTGTGAGCGGTGGAATCATCTTCATCATC 480
DB 474 GGGGCAATGAGAAGGTGAAGGCTCACAATCTGTGAGCGGTGGAATCATCTTCATCATC 533
QY 481 ACGGGATGGTGGTCTCATCTGTGAGCTGGGTGGCATGCGCATCATCAGAGATTC 540
DB 534 ACGGGATGGTGGTCTCATCTGTGAGCTGGGTGGCATGCGCATCATCAGAGATTC 593

QY 541 TATAACTCAATAGTAGTAATGTTGCCCAAAACCGTAGCTTTGGAGAGCTCTCTACTTAGGA 600
DB 594 TATTAACCAATAGTAGTAATGTTGCCCAAAACCGTAGCTTTGGAGAGCTCTCTACTTAGGA 653
QY 601 TGGACCAACGGCACTGGTGTCTGATTTGGAGAGCTCTGTTCTGCTGGTGTGTTTGTTC 660
DB 654 TGGACCAACGGCACTGGTGTCTGATTTGGAGAGCTCTGTTCTGCTGGTGTGTTTGTTC 713
QY 661 AAGCAAAAGAGCAGTAGTACAGATACCTTCCATCGCACCAACCCAAAAAGT 720
DB 714 AAGCAAAAGAGCAGTAGTACAGATACCTTCCATCGCACCAACCCAAAAAGT 773
QY 721 TATCACACCGGAAGAGTCAACCGAGCTTACTCAGAAAGTCAAGTATGTGTAGTTGTGT 780
DB 774 TATCACACCGGAAGAGTCA-----CGTCTACTCCAGAAAGTCAAGTATGTGTAGTTGTGT 828
QY 781 ATGTTTTTTTAACTTTACTATTAAGCCATGCAAAATGACAAAAATCTATATTCTTCTCA 840
DB 829 ATGTTTTTTTAACTTTACTATTAAGCCATGCAAAATGACAAAAATCTATATTCTTCTCA 888
QY 841 AAATGGACCCCAAGAAACCTTTGATTTACTGTTCTTAACTGCTTAATCTTAATTAAGGA 900
DB 889 AAATGGACCCCAAGAAACCTTTGATTTACTGTTCTTAACTGCTTAATCTTAATTAAGGA 948
QY 901 ACTGTGATCAGCTATTTATGATTTCTATTAAGCTTTTACGACAGATGAGATATTAACCC 960
DB 949 ACTGTGATCAGCTATTTATGATTTCTATTAAGCTTTTACGACAGATGAGATATTAACCC 1008
QY 961 AATGCTTTGATTTCTAGAAAGTATAGTAATTTGTTTTCTAAGTGGTTCAGACATCTA 1020
DB 1009 AATGCTTTGATTTCTAGAAAGTATAGTAATTTGTTTTCTAAGTGGTTCAGACATCTA 1068
QY 1021 CTCCTTTTATCATTTACTTCCAAAATGACATTTGCTTAAAGTGCATTTATTTTACTACTGTA 1080
DB 1069 CTCCTTTTATCATTTACTTCCAAAATGACATTTGCTTAAAGTGCATTTATTTTACTACTGTA 1128
QY 1081 ATTTCTCCAGACATAGCATTATGTATAGATAGTGTAACTATTTATATCTCACATAGA 1140
DB 1129 ATTTCTCCAGACATAGCATTATGTATAGATAGTGTAACTATTTATATCTCACATAGA 1188
QY 1141 GACATGCTTATATGTTTTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGA 1200
DB 1189 GACATGCTTATATGTTTTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGA 1248
QY 1201 ACTCAACTATTTGCTTTTTCAGGAAATCATGATAGGTTGAAGAGTGTACTATTAATTG 1260
DB 1249 ACTCAACTATTTGCTTTTTCAGGAAATCATGATAGGTTGAAGAGTGTACTATTAATTG 1308
QY 1261 TTTAAAAACAGCTTAGGGATTAAATGCTCTCCATTTTAAATGAAGATTAAATGAAGCTT 1320
DB 1309 TTTAAAAACAGCTTAGGGATTAAATGCTCTCCATTTTAAATGAAGATTAAATGAAGCTT 1368
QY 1321 TAATCAGCATTTGAAAGAAATTAAGTGGCTTCTGATATGCTGTTTTTAGCTTAGGAG 1380
DB 1369 TAATCAGCATTTGAAAGAAATTAAGTGGCTTCTGATATGCTGTTTTTAGCTTAGGAG 1428
QY 1381 TTAGAAATCCCTAACTCTTTTATCTCTCCAGAGGC- TTTTTTTTTTCTTCTGTATTA 1439
DB 1429 TTAGAAATCCCTAACTCTTTTATCTCTCCAGAGGC TTTTTTTTTTCTTCTGTATTA 1488
QY 1440 AATTAAACATTTTAAAAACGAGATATTTGTCAAGGGGCTTTGCAATTCAAACTGCTTTTC 1499
DB 1489 AATTAAACATTTTAAAAACGAGATATTTGTCAAGGGGCTTTGCAATTCAAACTGCTTTTC 1548
QY 1500 CAGGCTATCTCAGAGAGAGATAAAGTGTGATCTTAAGAAAAAGTGTGTTTTTAGGA 1559
DB 1549 CAGGCTATCTCAGAGAGAGATAAAGTGTGATCTTAAGAAAAAGTGTGTTTTTAGGA 1608
QY 1560 AAGTGAATAATTTTTTGTGTTTTTGAAGAGATGATGATTTTGAAGAAATCA 1619
DB 1609 AAGTGAATAATTTTTTGTGTTTTTGAAGAGATGATGATTTTGAAGAAATCA 1668
QY 1620 TATATGTATGATATATTTTAAATGATTTTGAAGTACAGACTTTGAGGTTTCATCAATAT 1679

442 GCTCACATTTCTGCTGAGCGGTGGAATCATCTTTCATCATCAGCGGCATGGTGGTCTCATC 501
Db GCTCACATTTCTGCTGAGCGGTGGAATCATCTTTCATCATCAGCGGCATGGTGGTCTCATC 480
QY CTTGTGAGCTGGTGGCAATGCCATCATCAGAGATTTCTATTAACATCAATAGTGAATGTT 561
Db CTTGTGAGCTGGTGGCAATGCCATCATCAGAGATTTCTATTAACATCAATAGTGAATGTT 540
QY GCCCAAAAAGCTGAGCTTGGAGAGCTCTCTACTTAGATGGACACCGCAGCTGGTGGTG 621
Db GCCCAAAAAGCTGAGCTTGGAGAGCTCTCTACTTAGATGGACACCGCAGCTGGTGGTG 600
QY ATTGTTGGAGAGCTCTGTTCTGCTGCTGTTTTTTTGTGCAACGAAAGAGCTAGCTAC 681
Db ATTGTTGGAGAGCTCTGTTCTGCTGCTGTTTTTTTGTGCAACGAAAGAGCTAGCTAC 660
QY AGATACCTCGATPACCTTCCCATCGCACAAACCAAAAAGTTTATCACACCAGCAAGAGTCA 741
Db AGATACCTCGATPACCTTCCCATCGCACAAACCAAAAAGTTTATCACACCAGCAAGAGTCA 720
QY CCGAGCGCTACTCCAGAGCTCAGTATGTGTAGTGTGTATGTTTTTTTAACTTACTAT 801
Db CCGAGCGCTACTCCAGAGCTCAGTATGTGTAGTGTGTATGTTTTTTTAACTTACTAT 780
QY AAAGCCATGCAAAATGACAAAATCTATATTACTTCTCAAAATGGACCCCAAGAACTT 861
Db AAAGCCATGCAAAATGACAAAATCTATATTACTTCTCAAAATGGACCCCAAGAACTT 840
QY TGAATTAAGCTTCTTAAGCTGCTTAATCTTAATTACAGAACTGTGCATCAGCTATTATG 921
Db TGAATTAAGCTTCTTAAGCTGCTTAATCTTAATTACAGAACTGTGCATCAGCTATTATG 900
QY ATTCTAATGCTATTTCAGGAAATGAGATTAATTAACCCAAATGCTTGTATGTTCTAGAA 981
Db ATTCTAATGCTATTTCAGGAAATGAGATTAATTAACCCAAATGCTTGTATGTTCTAGAA 960
QY AGTATAGTAATTTGTTTCTTAAGTGGTCTCAAGCATCTACTCTTTTATCATTTACTTCA 1041
Db AGTATAGTAATTTGTTTCTTAAGTGGTCTCAAGCATCTACTCTTTTATCATTTACTTCA 1020
QY AAATGACATTCCTAAAGCTGCAATTTATTTACTACTGTAATTTCTCCAGCATAGCAAT 1101
Db AAATGACATTCCTAAAGCTGCAATTTATTTACTACTGTAATTTCTCCAGCATAGCAAT 1080
QY ATGTACATAGATGTGTACATTTATATCTCATATAGACAGATGCTTATATGTTTTTAT 1161
Db ATGTACATAGATGTGTACATTTATATCTCATATAGACAGATGCTTATATGTTTTTAT 1140
QY TTAATAATGAAATGCCAGTCCCATTAACCTGAATAATAGAACTCAACTATTGCTTTTCAGG 1221
Db TTAATAATGAAATGCCAGTCCCATTAACCTGAATAATAGAACTCAACTATTGCTTTTCAGG 1200
QY GAAATCATGATGATGGTTGAAGAGGTTACTTATTAATGTTTAAACACAGCTTAGGGATT 1281
Db GAAATCATGATGATGGTTGAAGAGGTTACTTATTAATGTTTAAACACAGCTTAGGGATT 1260
QY AATGTCCTCCATTTATATGAAGATTAAATGAAGGCTTTTAATCAGCATTTGTAAGGAAA 1341
Db AATGTCCTCCATTTATATGAAGATTAAATGAAGGCTTTTAATCAGCATTTGTAAGGAAA 1320
QY TTGAATGCTTTCTGATATGCTGTTTTTTAGCCCTAGAGTTAGAAATCTCTTAATCTTTTA 1401
Db TTGAATGCTTTCTGATATGCTGTTTTTTAGCCCTAGAGTTAGAAATCTCTTAATCTTTTA 1380
QY TCCCTCTCTCCAGAGGCTTTTTTTTCTGCTGTATTAATTAACATTTTAAACGCGAG 1461
Db TCCCTCTCTCCAGAGGCTTTTTTTTCTGCTGTATTAATTAACATTTTAAACGCGAG 1440
QY ATATTTTCTCAAGGGCTTTTGCATTAACCTGCTTTTCCAGGGCTACTACTCAGAAGAAAG 1521
Db ATATTTTCTCAAGGGCTTTTGCATTAACCTGCTTTTCCAGGGCTACTACTCAGAAGAAAG 1500
QY ATAAAGTGTGATCTAGAAAAGTGTGTTTTAGGAAAAGTGAATAATTTTTTGTGTTTTT 1581

1501 ATAAAGTGTGATCTAAGAAAAGTGTGTTTTAGAAAAGTGAATAATTTTTTGTGTTTT 1560
QY GTATTTTGAAGAAGATGATGCAATTTTGAAGAAGATCATATATGATGATATATTTTAA 1641
Db GTATTTTGAAGAAGATGATGCAATTTTGAAGAAGATCATATATGATGATATATTTTAA 1620
QY TAAGTATTTGAGTACAGACTTTTGGAGTTCATCAATATATAAATAAGACAGCAAAAATA 1700
Db TAAGTATTTGAGTACAGACTTTTGGAGTTCATCAATATATAAATAAGACAGCAAAAATA 1679
RESULT 14
BC020866
LOCUS Homo sapiens claudin 8, mRNA (CDNA clone MGC:24067 IMAGE:4594155),
DEFINITION complete cds.
ACCESSION BC020866
VERSION BC020866.1 GI:18089189
SOURCE MGC.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 1835)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Faxmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
MEDLINE 12477932
PUBMED
2 (bases 1 to 1835)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
Center Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 36 Row: e Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21314655.

FEATURES		Location/Qualifiers	QY	
Source				
gene	1..1835	/organism="Homo sapiens"	601 TGGACCGGCACTGGTCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTC 660	
	/mol_type="mRNA"		614 TGGACCGGCACTGGTCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTC 673	
	/db_xref="taxon:9606"		661 AACGAAAGAGCAGTAGCTACAGATACCTTCGATACCTCCATCGCACAAACCCAAAAAGT 720	
	/clone="MGC:24067 IMAGE:4594155"		674 AACGAAAGAGCAGTAGCTACAGATACCTTCGATACCTCCATCGCACAAACCCAAAAAGT 733	
	/tissue_type="Kidney"		721 TATCACACCGGAAAGAGTCAACGAGGCTTACTCCAGAGTCACTGATGTGTGTGTGT 780	
	/clone_lib="NIH MGC_75"		734 TATCACACCGGAAAGAGTCAACGAGGCTTACTCCAGAGTCACTGATGTGTGTGTGT 793	
	/lab_host="DH10B"		781 ATGTTTTTTTTAACTTTACTATAAAGCCATGCAAAATGACAAAAATCTATATTCTTCTCA 840	
	/note="Vector: pDNR-LIB"		794 ATGTTTTTTTTAACTTTACTATAAAGC----- 819	
	1..1835		841 AATGGACCCCAAGAAACTTTGATTACTGTTCTTAACTGCTTAATCTTAATTACAGGA 900	
	/gene="CLDN8"		820 ----- 819	
CDS	/codon_start=1		901 ACTGTGCATCAGCTATTTATGATTTCTATAAGCTATTTTCAGCAGAATGAGATATTAAACCC 960	
	/product="claudin 8"		820 ----- 819	
	/protein_id="AAH20866.1"		961 AATGCTTTGATTGTTCTAGAAAGTATAGTAAATTTGTTTCTAAGGTGGTTCAAGCATCTA 1020	
	/db_xref="GI:18089190"		820 ----- 819	
	/db_xref="LocusID:9073"		1021 CTCCTTTTATCATTTCTTCAAAATGACATTTGCTTAAAGACTGCATTTATTTTACTACTGTA 1080	
	/translation="MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSAPIENNIIVFVE NFEGLMNCRQANIRWCKIYDLSLALSPDLQAARGKLMCAASVMSFLAFMMAALGM KCTRCTGNEKVKAHILLTAGIIFIITGMVLLIPSVWVANAIIRDYNSIIVNVAQKE LGEALYLVGTTALVLIVGGALPCCVFCNEKSSSYRYSIPSHETTKQSYHTGKKPSV YRSQYV"		820 -----CATTTACTTCAAAATGACATTTGCTTAAAGACTGCATTTATTTTACTACTGTA 869	
	194..655		1081 ATTTCTCCACGACATAGCATTTATGTACATAGATGAGTGAACATTTATATCTCACATAGA 1140	
	/note="PMP22_Claudin; Region: PMP-22/EMP/MP20/Claudin family"		870 ATTTCTCCACGACATAGCATTTATGTACATAGATGAGTGAACATTTATATCTCACATAGA 929	
	/db_xref="CDD:pfam00822"		1141 GACATGCTTATATGTTTTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGA 1200	
			930 GACATGCTTATATGTTTTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGA 989	
misc_feature	194..655		1201 ACTCAACTATGCTTTTCAGGGAATCATGATAGGTTGAAGAGTTACTATTAATTG 1260	
			990 ACTCAACTATGCTTTTCAGGGAATCATGATAGGTTGAAGAGTTACTATTAATTG 1049	
			1261 TTTAAAAACAGCTTAGGGATTAATGCTCTCCATTTATATGAAGATTAAAAATGAAGGCTT 1320	
			1050 TTTAAAAACAGCTTAGGGATTAATGCTCTCCATTTATATGAAGATTAAAAATGAAGGCTT 1109	
			1321 TAATCAGCATTTGTAAGGAAATGAAATGGCTTTCTGATATGCTGTTTTTTCAGCTTAGGAG 1380	
			1110 TAATCAGCATTTGTAAGGAAATGAAATGGCTTTCTGATATGCTGTTTTTTCAGCTTAGGAG 1169	
			1381 TTAGAATCCCTAACTTTTATCTCTCTCCAGAGGCTTTTTTTTTCTTCTGTGTATTAA 1440	
			1170 TTAGAATCCCTAACTTTTATCTCTCTCCAGAGGCTTTTTTTTTCTTCTGTGTATTAA 1229	
			1441 ATTAACATTTTTAAAAAGCAGATATTTTTGTCAAGGGGCTTTTGCAATCTCAAACTGCTTTCC 1500	
			1230 ATTAACATTTTTAAAAAGCAGATATTTTTGTCAAGGGGCTTTTGCAATCTCAAACTGCTTTCC 1289	
ORIGIN			1501 AGGGCTATCTCAGAGAAAGATATAAGTGTGATCTTAAGAAAAGTGAATGTTTTTAGGAA 1560	
			1290 AGGGCTATCTCAGAGAAAGATATAAGTGTGATCTTAAGAAAAGTGAATGTTTTTAGGAA 1349	
			1561 AGTGAATAATTTTTTGTGTTTTTGTATTTTGAAGAGAAATGATGATTTTTGACAAGAAATCAT 1620	
			1350 AGTGAATAATTTTTTGTGTTTTTGTATTTGAAGAGAAATGATGATTTTTGACAAGAAATCAT 1409	
			1621 ATATGATGAGATATATTTTAAAGTATTTTGAAGTACAGCTTTGAGGTTTTTCATCAATATA 1680	
			1410 ATATGATGAGATATATTTTAAAGTATTTTGAAGTACAGCTTTGAGGTTTTTCATCAATATA 1469	
Query Match	77.1%; Score 1550.4; DB 9; Length 1835;			
	Best Local Similarity 88.8%; Pred. No. 9.5e-301;			
	Matches 1785; Conservative 0; Mismatches 1; Indels 224; Gaps 1;			
QY	1 GGAATACTGTTCTTCTTCTGTCGACAGAGAACCTGCTTCAAAGCAGAGTAGCAGATTTC 60			
DB	14 GGAATACTGTTCTTCTTCTGTCGACAGAGAACCTGCTTCAAAGCAGAGTAGCAGATTTC 73			
QY	61 CGAGTCCAGCTGGCTTAAACTCATCCAGAGGATAATGGCAACCCATGCCCTTAGAAATC 120			
DB	74 CGAGTCCAGCTGGCTTAAACTCATCCAGAGGATAATGGCAACCCATGCCCTTAGAAATC 133			
QY	121 GCTGGCTGTTTTCTTGTGGTGTGGAATGTTGGGCACAGTGGCTGTCACTGTATGCTCT 180			
DB	134 GCTGGCTGTTTTCTTGTGGTGTGGAATGTTGGGCACAGTGGCTGTCACTGTATGCTCT 193			
QY	181 CAGTGAGAGTGTGGCTTCATTTGAAACACATCGTGGTTTTTGAAGACTTCTGGAA 240			
DB	194 CAGTGAGAGTGTGGCTTCATTTGAAACACATCGTGGTTTTTGAAGACTTCTGGAA 253			
QY	241 GCACTGTGGAATGCGTGGAGGAGCTTAACATCAGATGCAAGTCAAAATCTATGAT 300			
DB	254 GCACTGTGGAATGCGTGGAGGAGCTTAACATCAGATGCAAGTCAAAATCTATGAT 313			
QY	301 TCCTGCTGGCTTTTCTCCGACCTACAGCAGCCAGAGACTGATGTGTGCTGCTTC 360			
DB	314 TCCTGCTGGCTTTTCTCCGACCTACAGCAGCCAGAGACTGATGTGTGCTGCTTC 373			
QY	361 GTGATGTCCTTCTTGGCTTTTCAATGATGCCATCTTGGCATGAATGCAACAGGTGCAG 420			
DB	374 GTGATGTCCTTCTTGGCTTTTCAATGATGCCATCTTGGCATGAATGCAACAGGTGCAG 433			
QY	421 GGGGCAATGAGAGGTGAAGCTCACATTTCTGTCAGCGGTGGAATCATCTTCATCATC 480			
DB	434 GGGGCAATGAGAGGTGAAGCTCACATTTCTGTCAGCGGTGGAATCATCTTCATCATC 493			
QY	481 ACGGGCATGGTGTGCTCATCCCTGTGAGCTGGGTGGCAATGCCATCATCAGAGATTTC 540			
DB	494 ACGGGCATGGTGTGCTCATCCCTGTGAGCTGGGTGGCAATGCCATCATCAGAGATTTC 553			
QY	541 TATAACTCAATAGTGAATTTGCCCAAAAAAGTGAAGCTTTGGAGAGCTCTCTACTTAGGA 600			

QY	1741	1530	1801	1590	1861	1650	1921	1710
QY	AAAAAGATTTGCTCTTTGAGAACTTCACTGCTCCCTATGTGGTGACCTGAGTCAAAATTTGTC	1589	1801	1590	1861	1650	1921	1710
Db	1530	1801	1590	1861	1650	1921	1710	
QY	ATTTTGTGTTCTGTGAAAAATAAAATTTCCCTCTTGCTAGCAATTTCTGTTTGTAGTTTACTAAA	1860	1590	1861	1650	1921	1710	
Db	1590	1861	1650	1921	1710			
QY	ATCTGTAATACTGTGTAATTTTTCTGTTTATTTCCAAATTTGATGAACTGACAATCCCAATTT	1920	1590	1861	1650	1921	1710	
Db	1650	1921	1710					
QY	GAAAGTTTGTGCGACGCTCTGTAGCTTTAAATGAATGTGTTCTTATTTGCTTTTATACATT	1980	1590	1861	1650	1921	1710	
Db	1921	1710						
QY	GAAGTTTGTGCGACGCTCTGTAGCTTTAAATGAATGTGTTCTTATTTGCTTTTATACATT	1769	1590	1861	1650	1921	1710	
Db	1710							

1	10920	contig of	10920	bp in length
11021	18635	contig of	7615	bp in length
18736	27618	contig of	8883	bp in length
27719	33370	contig of	5652	bp in length
33471	38522	contig of	5052	bp in length
38623	44935	contig of	6313	bp in length
45036	52903	contig of	7868	bp in length
53004	59933	contig of	6930	bp in length
60034	55715	contig of	5682	bp in length
65816	72615	contig of	7800	bp in length
73716	79180	contig of	5465	bp in length
79281	83364	contig of	4084	bp in length
83465	87054	contig of	3590	bp in length
87155	91199	contig of	4045	bp in length

1	10920	contig of	10920	bp in length
11021	18635	contig of	7615	bp in length
18736	27618	contig of	8883	bp in length
27719	33370	contig of	5652	bp in length
33471	38522	contig of	5052	bp in length
38623	44935	contig of	6313	bp in length
45036	52903	contig of	7868	bp in length
53004	59933	contig of	6930	bp in length
60034	65715	contig of	5682	bp in length
65816	72615	contig of	7800	bp in length
73716	79180	contig of	5465	bp in length
79281	83364	contig of	4084	bp in length
83465	87054	contig of	3590	bp in length
87155	91199	contig of	4045	bp in length

1	10920	contig of	10920	bp in length
11021	18635	contig of	7615	bp in length
18736	27618	contig of	8883	bp in length
27719	33370	contig of	5652	bp in length
33471	38522	contig of	5052	bp in length
38623	44935	contig of	6313	bp in length
45036	52903	contig of	7868	bp in length
53004	59933	contig of	6930	bp in length
60034	65715	contig of	5682	bp in length
65816	72615	contig of	7800	bp in length
73716	79180	contig of	5465	bp in length
79281	83364	contig of	4084	bp in length
83465	87054	contig of	3590	bp in length
87155	91199	contig of	4045	bp in length

*	18636	18735: gap of 100 bp
*	18736	27618: contig of 8883 bp in length
*	27619	27718: gap of 100 bp
*	27719	33370: contig of 5652 bp in length
*	33371	33470: gap of 100 bp
*	33471	38522: contig of 5052 bp in length
*	38523	38822: gap of 100 bp
*	38823	44935: contig of 6313 bp in length
*	44936	45035: gap of 100 bp
*	45036	52903: contig of 7868 bp in length
*	52904	53003: gap of 100 bp
*	53004	59933: contig of 6930 bp in length
*	59934	60033: gap of 100 bp
*	60034	65715: contig of 5862 bp in length
*	65716	65815: gap of 100 bp
*	65816	73615: contig of 7800 bp in length
*	73616	73715: gap of 100 bp
*	73716	79180: contig of 5465 bp in length
*	79181	79280: gap of 100 bp
*	79281	83364: contig of 4084 bp in length
*	83365	83464: gap of 100 bp
*	83465	87054: contig of 3590 bp in length
*	87055	87154: gap of 100 bp
*	87155	91199: contig of 4045 bp in length
*	91200	91299: gap of 100 bp
*	91300	94966: contig of 3667 bp in length
*	94967	95066: gap of 100 bp
*	95067	99407: contig of 4341 bp in length
*	99408	99507: gap of 100 bp
*	99508	103203: contig of 3696 bp in length
*	103204	103303: gap of 100 bp
*	103304	106888: contig of 3585 bp in length
*	106889	106988: gap of 100 bp
*	106989	110187: contig of 3199 bp in length
*	110188	110287: gap of 100 bp
*	110288	113798: contig of 3511 bp in length
*	113799	113898: gap of 100 bp
*	113899	117663: contig of 3765 bp in length
*	117664	117763: gap of 100 bp
*	121135	121135: contig of 3372 bp in length
*	121136	121235: gap of 100 bp
*	121236	123224: contig of 1989 bp in length
*	123225	123324: gap of 100 bp
*	123325	125604: contig of 2280 bp in length
*	125605	125704: gap of 100 bp
*	125705	127990: contig of 2286 bp in length
*	127991	128090: gap of 100 bp
*	128091	130583: contig of 4593 bp in length
*	130584	130783: gap of 100 bp
*	130784	133371: contig of 2588 bp in length
*	133372	133471: gap of 100 bp
*	133472	136426: contig of 2955 bp in length
*	136427	136526: gap of 100 bp
*	136527	139256: contig of 2730 bp in length
*	139257	139356: gap of 100 bp
*	139357	141388: contig of 2032 bp in length
*	141389	141488: gap of 100 bp
*	141489	144089: contig of 2601 bp in length
*	144090	144189: gap of 100 bp
*	144190	146686: contig of 2497 bp in length
*	146687	146786: gap of 100 bp
*	146787	149873: contig of 3087 bp in length
*	149874	149973: gap of 100 bp
*	149974	151895: contig of 1922 bp in length
*	151896	151995: gap of 100 bp
*	151996	154840: contig of 2845 bp in length
*	154841	154940: gap of 100 bp
*	154941	157109: contig of 2169 bp in length
*	157110	157209: gap of 100 bp
*	157210	159650: contig of 2441 bp in length
*	159651	159750: gap of 100 bp
*	159751	161167: contig of 1417 bp in length
*	161168	161267: gap of 100 bp

	* 161268	163443: contig of 2176 bp in length	
	* 163444	165345: gap of 100 bp	
	* 163544	165635: contig of 2092 bp in length	
	* 165636	165735: gap of 100 bp	
	* 165736	168432: contig of 2697 bp in length	
	* 168433	168532: gap of 100 bp	
	* 168533	171331: contig of 2799 bp in length	
	* 171332	171431: gap of 100 bp	
Query Match	67.3%;	Score 1352.8; DB 2;	Length 191923;
Best Local Similarity	99.8%;	Pred. No. 2.9e-261;	
Matches 1365; Conservative	0;	Mismatches 2;	Indels 1; Gaps 1;
Qy	1	GGAAGAACTGTTCTTCTTCGTGGCACAGAGAACCCTGCTCCAAAGCAGAAAGTAGCAGTTTC	60
Dd	129314	GGAAGAACTGTTCTTCTTCGTGGCACAGAGAACCCTGCTTCGAAGCAGAAAGTAGCAGTTTC	129373
Qy	61	CGGAGTCCAAGTCGGCTPAAAAATCATCCACAGAGGATAATGGCAACCCATGCCTTAGAATC	120
Dd	129374	CGGAGTCCAAGTCGGCTPAAAAATCATCCACAGAGGATAATGGCAACCCATGCCTTA-AAATC	129432
Qy	121	GCTGGGCTGTTCTTGTTGGTGTTCGGAATGTTGGGCACAGTGGCTGTCACTGTCANGSCCT	180
Dd	129433	GCTGGGCTGTTCTTGTTGGTGTTCGGAATGTTGGGCACAGTGGCTGTCACTGTCAIGCCT	129492
Qy	181	CAGTGGAGAGTGTTCGCCCTTCATTGAAAACAACATCGTGGTTTTGAAAACTTCTCGGAA	240
Dd	129493	CAGTGGAGAGTGTTCGCCCTTCATTGAAAACAACATCGTGGTTTTGAAAACTTCTCGGAA	129552
Qy	241	GGACTGTGATGAAATTCGCTGAGGCAGGCTAACATCAGGATCGAGTCGAAATCTATGAT	300
Dd	129553	GGACTGTGATGAAATTCGCTGAGGCAGGCTAACATCAGGATCGAGTCGAAATCTATGAT	129612
Qy	301	TCCTCTGCTGGCTTTTCTCCGACCTCAGGCACCCAGAGGACTGATGTGTGCTTCTCC	360
Dd	129613	TCCTCTGCTGGCTTTTCTCCGACCTCAGGCACCCAGAGGACTGATGTGTGCTTCTCC	129672
Qy	361	GTGATGTCTTCTTGCTTTTCATGATGCCATCTCTGGCATGAAATGCACAGGTCCACG	420
Dd	129673	GTGATGTCTTCTTGCTTTTCATGATGCCATCTCTGGCATGAAATGCACAGGTCCACG	129732
Qy	421	GGGGACAATGAGAAGGTGAAGCTCACAATTCTGTGACGGCTGGAATCATCTTCATCATC	480
Dd	129733	GGGGACAATGAGAAGGTGAAGCTCACAATTCTGTGACGGCTGGAATCATCTTCATCATC	129792
Qy	481	ACGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGGCAATGCCATCATCAGAGATTTC	540
Dd	129793	ACGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGGCAATGCCATCATCAGAGATTTC	129852
Qy	541	TATAACTCAATAGTGAATGTGTGCCCAAACCGTGAGCTTGGAGAAGCTCTCTACTTTAGGA	600
Dd	129853	TATAACTCAATAGTGAATGTGTGCCCAAACCGTGAGCTTGGAGAAGCTCTCTACTTTAGGA	129912
Qy	601	TGGACCAACGGCATCTGGTGTCTGATTTGTGGAGAGCTCTGTCTGCTCGGTTTTTGTTCG	660
Dd	129913	TGGACCAACGGCATCTGGTGTCTGATTTGTGGAGAGCTCTGTCTGCTCGGTTTTTGTTCG	129972
Qy	661	AACGAAAGACGAGTAGCTACAGATACCTCGCATACCTTCCCACGCAACCCAAAAAGT	720
Dd	129973	AACGAAAGACGAGTAGCTACAGATACCTCGCATACCTTCCCACGCAACCCAAAAAGT	130032
Qy	721	TATCAACCCGAAAGAGTCAACGAGCGTCTATCCCAAGAGTCAGTATGTGTAGTTGTGT	780
Dd	130033	TATCAACCCGAAAGAGTCAACGAGCGTCTATCCCAAGAGTCAAGTATGTGTAGTTGTGT	130092
Qy	781	ATGTTTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAAATCTATATTCTTTCTCA	840
Dd	130093	ATGTTTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAAATCTATATTCTTTCTCA	130152
Qy	841	AAATGGAACCCCAAGAAACTTTTGATTTACTTTTAACTTGCCTTAATCTTAAATACAGA	900
Dd	130153	AAATGGAACCCCAAGAAACTTTTGATTTACTTTTAACTTGCCTTAATCTTAAATACAGA	130212

	* 161268	163443: contig of 2176 bp in length	
	* 163444	165345: gap of 100 bp	
	* 163544	165635: contig of 2092 bp in length	
	* 165636	165735: gap of 100 bp	
	* 165736	168432: contig of 2697 bp in length	
	* 168433	168532: gap of 100 bp	
	* 168533	171331: contig of 2799 bp in length	
	* 171332	171431: gap of 100 bp	
Query Match	67.3%;	Score 1352.8; DB 2;	Length 191923;
Best Local Similarity	99.8%;	Pred. No. 2.9e-261;	
Matches 1365; Conservative	0;	Mismatches 2;	Indels 1; Gaps 1;
Qy	1	GGAAGAACTGTTCTTCTTCGTGGCACAGAGAACCCTGCTCCAAAGCAGAAAGTAGCAGTTTC	60
Dd	129314	GGAAGAACTGTTCTTCTTCGTGGCACAGAGAACCCCTGCTTCAAGCAGAAAGTAGCAGTTTC	129373
Qy	61	CGGAGTCCAAGTCGGCTPAAAAATCATCCACAGAGGATAATGGCAACCCATGCCTTAGAATC	120
Dd	129374	CGGAGTCCAAGTCGGCTPAAAAATCATCCACAGAGGATAATGGCAACCCATGCCTTA-AAATC	129432
Qy	121	GCTGGGCTGTTTCTTGTTGGTGTTCGGAATGTTGGGCACAGTGGCTGTCACTGTCAAGCCCT	180
Dd	129433	GCTGGGCTGTTTCTTGTTGGTGTTCGGAATGTTGGGCACAGTGGCTGTCACTGTCAAGCCCT	129492
Qy	181	CAGTGGAGAGTGTTCGCCCTTCATTGAAAACAACATCGTGGTTTTGAAAACTTCTCGGAA	240
Dd	129493	CAGTGGAGAGTGTTCGCCCTTCATTGAAAACAACATCGTGGTTTTGAAAACTTCTCGGAA	129552
Qy	241	GGACTGTGATGAAATTCGCTGAGGCAGGCTTAACATCAGGATCAGTGCRAAAATCTATGAT	300
Dd	129553	GGACTGTGATGAAATTCGCTGAGGCAGGCTTAACATCAGGATCAGTGCRAAAATCTATGAT	129612
Qy	301	TCCTCTGCTGGCTTTTCTCCGACACTCAGGCACCCAGAGGACTGATGTGTGCTTCTCC	360
Dd	129613	TCCTCTGCTGGCTTTTCTCCGACACTCAGGCACCCAGAGGACTGATGTGTGCTTCTCC	129672
Qy	361	GTGATGTCTTCTTGGCTTTTATGATGCCATCTTGGCATGAAATGCACAGGTCCACG	420
Dd	129673	GTGATGTCTTCTTGGCTTTTATGATGCCATCTTGGCATGAAATGCACAGGTCCACG	129732
Qy	421	GGGGACAATGAGAAGGTGAAGCTCACAATTCTGTGACGGCTGGAATCATCTTTCATCATC	480
Dd	129733	GGGGACAATGAGAAGGTGAAGCTCACAATTCTGTGACGGCTGGAATCATCTTTCATCATC	129792
Qy	481	ACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTCGCAATGCCATCATCAGAGATTTC	540
Dd	129793	ACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTCGCAATGCCATCATCAGAGATTTC	129852
Qy	541	TATAATCTCAATAGTGAATGTGTGCCCAAACCGTGAGCTTGGAGAAGCTCTCTACTTTAGGA	600
Dd	129853	TATAATCTCAATAGTGAATGTGTGCCCAAACCGTGAGCTTGGAGAAGCTCTCTACTTTAGGA	129912
Qy	601	TGGACCAACGGCATCTGGTGTCTGATTTGTGGAGAGCTCTGTTCGTGCGTTTTTGTTCG	660
Dd	129913	TGGACCAACGGCATCTGGTGTCTGATTTGTGGAGAGCTCTGTTCGTGCGTTTTTGTTCG	129972
Qy	661	AACGAAAGACGAGTAGCTACAGATACCTCGCATACCTTCCCATCGCACACCCAAAAAGT	720
Dd	129973	AACGAAAGACGAGTAGCTACAGATACCTCGCATACCTTCCCATCGCACACCCAAAAAGT	130032
Qy	721	TATCAACCCGAAAGAGTACCGAGCGTCTATCCGAGAGTCAGTATGTGTAGTTGTGT	780
Dd	130033	TATCAACCCGAAAGAGTACCGAGCGTCTATCCGAGAGTACAGTATGTGTAGTTGTGT	130092
Qy	781	ATGTTTTTTTAACTTTTACTATTAAGCCATGCAAAATGACAAAAATCTATATCTTTCTCA	840
Dd	130093	ATGTTTTTTTAACTTTTACTATTAAGCCATGCAAAATGACAAAAATCTATATCTTTCTCA	130152
Qy	841	AAATGGAACCCCAAGAAACTTTTGAATTTACTTTTAACTTGCCTTAATCTTAAATACAGA	900
Dd	130153	AAATGGAACCCCAAGAAACTTTTGAATTTACTTTTAACTTGCCTTAATCTTAAATACAGA	130212

QY	901	ACTGTCATCAGCTATTATGATTTCTATAGCTATTTCAGCAGATGAGATATTAAACC	960
Db	130213	ACTGTCATCAGCTATTATGATTTCTATAGCTATTTCAGCAGATGAGATATTAAACC	130272
QY	961	AATGCTTTGATTTGTTCTAGAAAGTAGTAAATTTGTTTCTAAGGTGGTTCAAGCATCTA	1020
Db	130273	AATGCTTTGATTTGTTCTAGAAAGTAGTAAATTTGTTTCTAAGGTGGTTCAAGCATCTA	130332
QY	1021	CTCTTTTATCAATTTACTTCAAAATGACATGCTTAAGACTGCATTTATTTACTACTGA	1080
Db	130333	CTCTTTTATCAATTTACTTCAAAATGACATGCTTAAGACTGCATTTATTTACTACTGA	130392
QY	1081	ATTTCTCCACGACATAGCATTTATGTACATAGATGAGTAAACATTTATATCTCACATAGA	1140
Db	130393	ATTTCTCCACGACATAGCATTTATGTACATAGATGAGTAAACATTTATATCTCACATAGA	130452
QY	1141	GACATGCTTATATGTTTATTTAAATGAAATGCCAGTCCATTCACACTGAATAAATAGA	1200
Db	130453	GACATGCTTATATGTTTATTTAAATGAAATGCCAGTCCATTCACACTGAATAAATAGA	130512
QY	1201	ACTCAACTATTGCTTTTCAGGGAATCATGGATAGAGGTTTGAAGAAGTTTACTATTAAATTG	1260
Db	130513	ACTCAACTATTGCTTTTCAGGGAATCATGGATAGAGGTTTGAAGAAGTTTACTATTAAATTG	130572
QY	1261	TTTAAACACAGCTTAGGATTAATGCTCCTCATTTATATGAAGATTAATAAGAGGCTT	1320
Db	130573	TTTAAACACAGCTTAGGATTAATGCTCCTCATTTATATGAAGATTAATAAGAGGCTT	130632
QY	1321	TAATCAGCATTTGAAGGAATTCGAATGGCTTTCTGATATGCTGTTTT	1368
Db	130633	TAATCAGCATTTGAAGGAATTCGAATGGCTTTCTGATATGCTGTTTT	130680

Search completed: September 2, 2004, 23:32:10
Job time : 5296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 22:04:00 ; Search time 3550 Seconds
(without alignments)
16907.883 Million cell updates/sec

Title: US-10-063-732-119

Perfect score: 2010

Sequence: 1 ggaaaaactgtctctctctg.....aattgtacattttcttaatt 2010

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_nam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rpd:**

26: em_gss_rpd:**

27: em_gss_vrt:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	761	37.9	837	9 AU121779
2	719	35.8	754	12 BG674625
3	716	35.6	772	12 BG426438
4	683.4	34.0	703	14 CB851106

5	678	33.7	678	29	AY400665	Homo sapi
6	587.6	29.2	903	12	BG402103	
7	552	27.5	559	10	AW235670	xn21c12.x
8	542.4	27.0	544	10	BF195920	7087c12.x
9	541	26.9	547	10	BF196092	hr8ah10.x
10	535.8	26.7	577	9	AU147533	AU147533
11	525.2	26.1	551	14	CB851224	UI-CF-DU1
12	514.4	25.6	586	12	BG957666	PMO-CT080
13	512	25.5	513	10	AW237740	xm81a03.x
14	505.2	25.1	689	10	BF675808	602083725
15	494	24.6	678	29	AY400667	Mus muscu
16	480.4	23.9	492	10	AW205135	UI-H-B11-
17	468.4	23.3	683	14	CB418866	591692 MA
18	462.2	23.0	681	14	CF766535	CES003314
19	460.2	22.9	633	14	CF169716	B0817B10-
20	445.2	22.1	809	12	BG402412	602456111
21	443.2	22.0	902	12	BI852250	603378454
22	440.2	21.9	557	10	BF231401	253787 BA
23	438.8	21.8	486	14	CA313163	UI-CF-FNO
24	429	21.3	688	14	CF796030	892243 MA
25	427.8	21.3	693	14	CF169015	B0808A07-
26	425	21.0	490	9	AA536178	nf96b03.s
27	421.6	21.0	444	14	CF140548	UI-HF-C80
28	404.4	20.1	752	9	AJ507044	AJ507044
29	399	19.9	400	10	AW237781	xm81807.x
30	398.6	19.8	642	14	CK333041	H8225C06-
31	397.4	19.8	656	14	CF168607	B0802A03-
32	396.6	19.7	551	9	AI786907	u1j2g11.y
33	390.8	19.4	754	12	BI104115	602890432
34	365.8	18.2	580	14	CB455630	712787 MA
35	360.8	18.0	584	14	CB435534	600565 MA
36	360.4	17.9	918	10	BE285566	601035576
37	355.6	17.7	767	14	CB235084	AGENCOURT
38	355	17.7	355	13	EX283446	EX283446
39	351	17.5	360	10	BE172635	MRO-HT055
40	347.2	17.3	352	10	BE000282	MRO-BN007
41	345.2	17.2	351	10	BE814313	MRO-BN007
42	339.8	16.9	596	10	BE618773	BB618773
43	336	16.7	363	14	CB306981	UI-CF-FNO
44	332.4	16.5	568	14	CA339169	NISC_lx11
45	330.6	16.4	465	9	AI048929	uc84g08.y
46	328.2	16.3	771	9	AI547146	PN3.1.01
47	325.2	16.2	643	10	BB610589	BB610589
48	323.4	16.1	325	29	AY400666	Pan trogl
49	319.8	15.9	578	10	BF151294	u214804.y
50	311.2	15.5	348	9	AA769736	OB20C04.s
51	311.2	15.5	375	14	CF140477	UI-HF-C80
52	284.4	14.1	459	12	BI403157	MI-P-CPI-
53	284.2	14.1	626	14	CF171768	B0847B12-
54	282.6	14.1	500	10	BE483534	169725 BA
55	278.8	13.9	862	13	EX720601	EX720601
56	278.2	13.8	482	13	EQ285325	NISC_ff05
57	277.6	13.8	675	29	AY401721	Homo sapi
58	276.4	13.8	613	14	CF171518	B0843E03-
59	271.2	13.5	675	29	AY401722	Pan trogl
60	270.6	13.5	1174	11	AK048287	Mus muscu
61	269.4	13.4	856	14	CB236303	AGENCOURT
62	267.8	13.3	675	29	AY401723	Mus muscu
63	256.8	12.8	268	10	AW999412	MRO-BN007
64	251.2	12.5	256	10	BE172270	MRO-H1055
65	245.2	12.2	465	9	AI037192	ud60ail.y
66	243.8	12.1	642	14	CB457008	714331 MA
67	241.6	12.0	648	28	BZ846477	CH240.193
68	239.4	11.9	316	10	BE476649	160002 BA
69	233.8	11.6	266	10	BF329750	RCC-BN028
70	233.6	11.6	593	14	CB433030	609301 MA
71	233.2	11.6	441	9	AA839199	vw93h03.z
72	225	11.2	682	12	BI401450	MI-P-CPO-
73	221.2	11.0	397	10	AW990108	ms65h08.y
74	221	11.0	521	10	AW990108	uf34501.y
75	219.8	10.9	402	9	AA167924	ms65h08.r
76	217	10.8	585	14	CF114057	Shulzeoni
77	215.4	10.7	535	10	AW226871	um62912.y

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10594 row: c column: 19
High quality sequence stop: 750.
Location/Qualifiers
1..754
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4746186"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 35.8%; Score 719; DB 12; Length 754;
Best Local Similarity 99.6%; Pred. No. 1e-133;
Matches 752; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 471 CTTTCATCATCGGCGATGCTGCTCTCATCTGCTGAGCTGGTGGCAATGCCATCAT 530
DB 1 CTTTCATCATCGGCGATGCTGCTCTCATCTGCTGAGCTGGTGGCAATGCCATCAT 60
QY 531 CAGAGATTCTTATACTCAATAGTGAATGTTGCCAAAACCTGAGCTTGGAGAAGCTCT 590
DB 61 CAGAGATTCTTATACTCAATAGTGAATGTTGCCAAAACCTGAGCTTGGAGAAGCTCT 120
QY 591 CTACTTAGATCGACACCGGCACTGGTCTGATTTGGAGAGCTCTG-TTCTGCTGGC 649
DB 121 CTACTTAGATCGACACCGGCACTGGTCTGATTTGGAGAGCTCTGTTCTGCTGGC 180
QY 650 TTTTGTGTGCAACGAAGAAGACGCTAGCTACAGATCTCCATACCTTCCCATCGCAAA 709
DB 181 TTTTGTGTGCAACGAAGAAGACGCTAGCTACAGATCTCCATACCTTCCCATCGCAAA 240
QY 710 CCCAAAAAGTTATCACACCGGAAAGAGTCCAGAGCTCTACTCCAGAGTCAAGTATG 769
DB 241 CCCAAAAAGTTATCACACCGGAAAGAGTCCAGAGCTCTACTCCAGAGTCAAGTATG 300
QY 770 TGTAGTTGTATGTTTTTTTAACTTTACTATAAGCCATGCAATGACAAAATCTATA 829
DB 301 TGTAGTTGTATGTTTTTTTAACTTTACTATAAGCCATGCAATGACAAAATCTATA 360
QY 830 TTACTTTCTCAAAATGGACCCCAAGAAACTTGTATTTACTGTTCTTAACTGCCATACT 889
DB 361 TTACTTTCTCAAAATGGACCCCAAGAAACTTGTATTTACTGTTCTTAACTGCCATACT 420
QY 890 TAATTACAGAACTGTGCATAGCTATTATGATTTCTATAGCTATTTCACGAGAATGAG 949
DB 421 TAATTACAGAACTGTGCATAGCTATTATGATTTCTATAGCTATTTCACGAGAATGAG 480
QY 950 ATATTAAACCCCAATGTTTGTGTTAGAAAGPATAGTAAATTTGTTTCTAAGGTGGT 1009
DB 481 ATATTAAACCCCAATGTTTGTGTTAGAAAGPATAGTAAATTTGTTTCTAAGGTGGT 540
QY 1010 TCAAGCATCTACTCTTTTATCATTTTCTCAAAATGACATTTGCTAAAGACTGCATTATT 1069
DB 541 TCAAGCATCTACTCTTTTATCATTTTCTCAAAATGACATTTGCTAAAGACTGCATTATT 600
QY 1070 TTTACTACTGTAATTTCTCCACACATAGCAATTTATGATAGTAGAGTGT-AACTTTAT 1128
DB 601 TTTACTACTGTAATTTCTCCACACATAGCAATTTATGATAGTAGAGTGTAAACATTAT 660
QY 1129 ATCTCATAAGACATGCTTATATGTTTATTTTAAATGAAATGCCAGTCCATTACAC 1188

DB 661 ATCTCATAAGACATGCTTATATGTTTTTATTAATAAATGCCATCCATTACAC 720
QY 1189 TGAATAAATAAGTCAATCAATATTCCTTTTCAGGGA 1223
DB 721 TGAATAAATAAGTCAATCAATATTCCTTTTCAGGGA 754

RESULT 3
BG426438 772 bp mRNA linear EST 14-MAR-2001
LOCUS 602492836F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4606880 5',
DEFINITION mRNA sequence.
ACCESSION BG426438
VERSION BG426438.1 GI:13332944
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1346 row: g column: 09
High quality sequence stop: 735.
Location/Qualifiers
1..772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4606880"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCGCATATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 35.6%; Score 716; DB 12; Length 772;
Best Local Similarity 99.3%; Pred. No. 4.1e-133;
Matches 760; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 25 ACAGAGAACCTCTGCTTCAAGACAGAGTAGCAGTCCGAGTCCAGCTGCTGCTAAATCTCA 84
DB 1 ACAGAGAACCTCTGCTTCAAGACAGAGTAGCAGTCCGAGTCCAGCTGCTGCTAAATCTCA 60
QY 85 TCCAGAGGATTAATGCAACCCATGCCCTTAGAATCGCTGGCTGTTCTTGGTGGTGT 144
DB 61 TCCAGAGGATTAATGCAACCCATGCCCTTAGAATCGCTGGCTGTTCTTGGTGGTGT 120
QY 145 GGAATGTTGGGACAGTGGCTGTCTACTGTCTCATGCTCAGTGGAGAGTGTGGCTTCATT 204
DB 121 GGAATGTTGGGACAGTGGCTGTCTACTGTCTCATGCTCAGTGGAGAGTGTGGCTTCATT 180
QY 205 GAAACACATCGTGGTTTTTGAAACCTTCGGAGAGACTGTGGATGAATTCGCTGAGG 264

Seq primer: M13 FORWARD			
POLYA=yes			
Location/Qualifiers			
1..703			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="UI-CF-EN1-adg-i-18-0-UI"			
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"			
/dev_stage="Adult"			
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"			
/clone_lib="UI-CF-EN1"			
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; the UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.			
TAG_SEQ=None found			
ORIGIN			
Query Match 34.0%; Score 683.4; DB 14; Length 703;			
Best local Similarity 99.9%; Pred. No. 1.4e-126;			
Matches 684; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1326	AGCATTTGAAGGAATTCGATGGCTTCTGATATGCTTTTATAGCTAGGAGTTAGA	1385
Db	703	AGCATTTGAAGGAATTCGATGGCTTCTGATATGCTTTTATAGCTAGGAGTTAGA	644
Qy	1386	AATCCTAACTCTTTATCCTCTTCCAGAGGCTTTTCTTGTGTATTAATAA	1445
Db	643	AATCCTAACTCTTTATCCTCTTCCAGAGGCTTTTCTTGTGTATTAATAA	584
Qy	1446	CATTTTAAACGAGATATTTTGTCAAGGGCTTTCATCAACATGCTTTCCAGGGC	1505
Db	583	CATTTTAAACGAGATATTTTGTCAAGGGCTTTCATCAACATGCTTTCCAGGGC	524
Qy	1506	TATACTCAGAGAAGATATAAGTGTGATCTAAGAAAAAGTGTGTTTATAGGAAGTGA	1565
Db	523	TATACTCAGAGAAGATATAAGTGTGATCTAAGAAAAAGTGTGTTTATAGGAAGTGA	464
Qy	1566	AAATATTTTGTGTTTATTTGAAGAGATGATGATTTTGAAGAAGATCATATATG	1625
Db	463	AAATATTTTGTGTTTATTTGAAGAAGATGATGATTTTGAAGAAGATCATATATG	404
Qy	1626	TATGGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGTTTCATCAATATAA	1685
Db	403	TATGGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGTTTCATCAATATAA	344
Qy	1686	AAGAGCAGAAAAATATGCTTGGTTTTCATTTGCTTACAAAAAACAAACAAAAA	1745
Db	343	AAGAGCAGAAAAATATGCTTGGTTTTCATTTGCTTACAAAAAACAAACAAAAA	284
Qy	1746	GTGTCTCTTGAAGACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATGTGCT	1805
Db	283	GTGTCTCTTGAAGACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATGTGCT	224
Qy	1806	TGTTCTGTGAAAAATAAATTTCCCTTCTGTACCAATTTCTGTTAGTTTACTAAATCTG	1865
Db	223	TGTTCTGTGAAAAATAAATTTCCCTTCTGTACCAATTTCTGTTAGTTTACTAAATCTG	164
Qy	1866	TAAATCTGATTTTCTGTTTATTTCCAAATTTGATGAACCTGCAATCCCAATTTGAAG	1925
Db	163	TAAATCTGATTTTCTGTTTATTTCCAAATTTGATGAACCTGCAATCCCAATTTGAAG	104

Query Match	33.7%	Score 678;	DB 29;	Length 678;
Best Local Similarity	100.0%;	Pred. No. 1.7e-125;		
Matches 678;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	97	ATGGCAACCCATGCGCTTAGAAATCGCTGGGCTCTTTCTTGTCGGTGTGGAAATGTGGGC	158	
Db	1	ATGGCAACCCATGCGCTTAGAAATCGCTGGGCTCTTTCTTGTCGGTGTGGAAATGTGGGC	60	
QY	157	ACAGTGGCTGCATCTGCATCGCTCAGTGGAGAGTGTGCGCCTTCATTGAAACAAACATC	216	
Db	61	ACAGTGGCTGCATCTGCATCGCTCAGTGGAGAGTGTGCGCCTTCATTGAAACAAACATC	120	
QY	217	GTGGTTTTTGAACCTTCCTGGGAAGACTGTGGATGAATTCGTGAGGAGGCTTAACATC	276	
Db	121	GTGGTTTTTGAACCTTCCTGGGAAGACTGTGGATGAATTCGTGAGGAGGCTTAACATC	180	
QY	277	AGGATGCAGTGCACAAATCTATGATCCCTCTCGCTCTTTCTCCGAGACTACAGCAGCC	336	
Db	181	AGGATGCAGTGCACAAATCTATGATCCCTCTCGCTCTTTCTCCGAGACTACAGCAGCC	240	
QY	337	AGAGGACTGATGTGTCGTCCTTCGTGATGTCCTTCTTGGCTTTGATGATGCCATCCTT	396	

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGC Library."

ORIGIN

Query Match	29.2%;	Score 587.6;	DB 12;	Length 903;
Best Local Similarity	99.0%;	Pred. No. 2.1e-107;		
Matches 612;	Conservative 0;	Mismatches 4;	Indels 2;	Gaps 2;
QY	242	GACTGTGATGAATTTCCGTGAGCAGGCTAACATCAGGATCAGGATGCAGTGCAGTGCAGGATCTATGATT	301	
DB	1	GACTGTGGATGAAATGGTGGAGCAGGCTAACATCAGGATCAGGATGCAGTGCAGGATCTATGATT	60	
QY	302	CCCTGCTGGCTCTTTCTCCGGACCTACAGCAGCCAGAGACTGATGTGCTGCTTCGG	361	
DB	61	CCCTGCTGGCTCTTTCTCCGGACCTACAGCAGCCAGAGACTGATGTGCTGCTTCGG	120	
QY	362	TGATGTCTCTTTGGCTTTTCATGATGGCCATCTCTTGGCATGAAATGACCAAGTGCACGG	421	
DB	121	TGATGTCTCTTTGGCTTTTCATGATGGCCATCTCTTGGCATGAAATGACCAAGTGCACGG	180	
QY	422	GGGACAAATGAGAAGGTGAAGGCTCACATTTCTGTGACGGCTGGAATCATCTTTCATCATCA	481	
DB	181	GGGACAAATGAGAAGGTGAAGGCTCACATTTCTGTGACGGCTGGAATCATCTTTCATCATCA	240	
QY	482	CGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCCT	541	
DB	241	CGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCCT	300	
QY	542	ATAACTCAATATGATGAATTTGCCCAAAACGCTGAGCTTGGAGAAGCTCTCTACTTAGGAT	601	
DB	301	ATAACTCAATATGATGAATTTGCCCAAAACGCTGAGCTTGGAGAAGCTCTCTACTTAGGAT	360	
QY	602	GGACACCGGCACTGGTGTGTGATTTGGAGAGAGCTCTGTTCTGCTGCGTTTGTGTTGCA	661	
DB	361	GGACACCGGCACTGGTGTGTGATTTGGAGAGAGCTCTGTTCTGCTGCGTTTGTGTTGCA	420	
QY	662	ACGAAAAGAGCAGTAGCTTACAGATPACTCGATACCTTCCCATCGCACACCCAAAAAGTT	721	
DB	421	ACGAAAAGAGCAGTAGCTTACAGATPACTCGATACCTTCCCATCGCACACCCAAAAAGTT	480	
QY	722	ATCACACCGGAAAAGAGTACCGAGCGTCTTACTTCCAGAAAGTCAGATGTGTAGTTGTGTA	781	
DB	481	ATCACACCGGAAAAGAGTACCGAGCGTCTTACTTCCAGAAAGTCAGATGTGTAGTTGTGTA	540	
QY	782	TGTTTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAATCTATATTAC-TTTTCCTCA	840	
DB	541	TG-TTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAATCTATATTACTTTTCTCA	599	
QY	841	AAATGGACCCCAAGAAA	858	
DB	600	AAATGGACCCCAAGAAA	617	

RESULT 7	AW235670/c	559 bp	linear	EST 13-DEC-1999
LOCUS	xn2ic12.x1			
DEFINITION	NCI CGAP Kid11 Homo sapiens cdna clone IMAGE:2694358 3', mRNA sequence.			

ACCESSION	AW235670	
VERSION	AW235670.1	GI:6568059
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	
	Eukaryota;	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia;	Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1	(bases 1 to 559)
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov	

REMARKS

source

1. .559

1. .559

1. .559

1. .559

1. .559

1. .559

1. .559

1. .559

1. .559

ORIGIN

Query Match	27.5%	Score 552	DB 10	Length 559
Best Local Similarity	100.0%	Pred. No. 3.1e-100		
Matches 552	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1459	CAGATATTTTGTCAAGGGGCTTTGCGATTCAAACTGCTTTCCAGGGCTATCTCAGAAGA	1518	
Db	559	CAGATATTTTGTCAAGGGGCTTTGCGATTCAAACTGCTTTCCAGGGCTATCTCAGAAGA	500	
Qy	1519	AGATAAAGTGTGATCTAGAAAAAGTCATGGTTTTAGGAAAGTGAATAATTTTTTGT	1578	
Db	499	AGATAAAGTGTGATCTAGAAAAAGTCATGGTTTTAGGAAAGTGAATAATTTTTTGT	440	
Qy	1579	TTTGTATTGGAAGAAGATGATGCATTTTGCACAAAGAAATCATATATGATGGATATATTT	1638	
Db	439	TTTGTATTGGAAGAAGATGATGCATTTTGCACAAAGAAATCATATATGATGGATATATTT	380	
Qy	1639	TAATAAGTATTTGAGTACAGACITTTGAGGTTTTCATCAATATAAAATAAAGAGCAGAAAAA	1698	
Db	379	TAATAAGTATTTGAGTACAGACITTTGAGGTTTTCATCAATATAAATAAAGAGCAGAAAAA	320	
Qy	1699	TATGTCCTGGTTTTTCATTTGCTTACAAAAAACAACAAGTTGTCCTTTGAG	1758	
Db	319	TATGTCCTGGTTTTTCATTTGCTTACAAAAAACAACAAGTTGTCCTTTGAG	260	
Qy	1759	AACTTCACCTGCTCCTATGTGGGTGACCTCGAGTCAAAATGTGCATTTTGTCTCTGAAAA	1818	
Db	259	AACTTCACCTGCTCCTATGTGGGTGACCTCGAGTCAAAATGTGCATTTTGTCTCTGAAAA	200	
Qy	1819	ATAAATTTCCCTCTGTACCAATTTCTGTTTAGTTTTACTAAATCTGTAATCTGTATT	1878	
Db	199	ATAAATTTCCCTCTGTACCAATTTCTGTTTAGTTTTACTAAATCTGTAATCTGTATT	140	
Qy	1879	TTTCTGTTTATTCCAAATTTTGATGAAACTGACAATCCAAATTTGAAAGTTTGTGCGACGT	1938	
Db	139	TTTCTGTTTATTCCAAATTTTGATGAAACTGACAATCCAAATTTGAAAGTTTGTGCGACGT	80	
Qy	1939	CTGCTAGCTTAATGAATGCTGTCATTTGCTTTATACATTTATATTAATAAATCTGAC	1998	
Db	79	CTGCTAGCTTAATGAATGCTGTCATTTGCTTTATACATTTATATTAATAAATCTGAC	20	
Qy	1999	ATTTCCTCAATT	2010	
Db	19	ATTTCCTCAATT	8	


```

RESULT 8
BF195920/c
LOCUS
DEFINITION
7087c12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3643390 3',
mRNA sequence.
ACCESSION
BF195920
VERSION
BF195920.1 GI:11083309
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 544)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
High quality sequence stop: 467.
FEATURES
Location/Qualifiers
1..544
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3643390"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not 1; Site: 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP Kid1 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
ORIGIN
Query Match 27.0%; Score 542.4; DB 10; Length 544;
Best Local Similarity 99.8%; Pred. No. 2.6e-98;
Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1064 ATATTATTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGAACA 1123
DB 544 ATATTATTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGAACA 485
QY 1124 TTATATCTCACATAGAGACATGCTTTATATGTTTATTTAAATGAATGCCAGTCCAT 1183
DB 484 TTTATATCTCACATAGAGACATGCTTTATATGTTTATTTAAATGAATGCCAGTCCAT 425
QY 1184 TACACTGAATAAATAGAACTCAACTATTCCTTTTCAGGGAATCATGATAGGGTTGAAG 1243
DB 424 TACACTGAATAAATAGAACTCAACTATTCCTTTTCAGGGAATCATGATAGGGTTGAAG 365
QY 1244 AAGGTTTACTATTAATTTGTTTAAACACAGGTTAGGATTAAATGCTCCATTATAATGAA 1303
DB 364 AAGGTTTACTATTAATTTGTTTAAACACAGGTTAGGATTAAATGCTCCATTATAATGAA 305
QY 1304 GATTAATAATGAGCGCTTAAATGAGCATTTCAAGGAATTTGATGCTTCTGATATGCT 1363
DB 304 GATTAATAATGAGCGCTTAAATGAGCATTTCAAGGAATTTGATGCTTCTGATATGCT 245

```

```

QY 1364 GTTTTTCAGCTAGGAGTTAGAAATCCTAACTCTTTATCTCTCTCCAGAGGCTTTT 1423
DB 244 GTTTTTCAGCTAGGAGTTAGAAATCCTAACTCTTTATCTCTCTCCAGAGGCTTTT 185
QY 1424 TTTTCTTGTGTATTAATAACATTTTAAACGCGAGATATTTTGTCAAGGGGCTTTGC 1483
DB 184 TTTTCTTGTGTATTAATAACATTTTAAACGCGAGATATTTTGTCAAGGGGCTTTGC 125
QY 1484 ATTCAAACTGCTTTTCCAGGGCTATCTACAGAGAAAGATATAAAGTGTGATCTAAGAAA 1543
DB 124 ATTCAAACTGCTTTTCCAGGGCTATCTACAGAGAAAGATATAAAGTGTGATCTAAGAAA 65
QY 1544 AGTGATGGTTTTAGAAAGTGAATAATTTTGTGTTTGTATTTGAAGAGAAATGATGCA 1603
DB 64 AGTGATGGTTTTAGAAAGTGAATAATTTTGTGTTTGTATTTGAAGAGAAATGATGCA 5
QY 1604 TTTT 1607
DB 4 TTTT 1
RESULT 9
BF196092/c
LOCUS
DEFINITION
hr8h10.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3134947 3',
mRNA sequence.
ACCESSION
BF196092
VERSION
BF196092.1 GI:11083657
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 547)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
High quality sequence stop: 472.
FEATURES
Location/Qualifiers
1..547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3134947"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site: 1: Not 1; Site: 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
ORIGIN
Query Match 26.9%; Score 541; DB 10; Length 547;
Best Local Similarity 100.0%; Pred. No. 4.9e-98;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1470 TCAGGGGCTTGGCTTCAAACTGCTTTTCCAGGCTATCTACTCAGAGAAAGATAAAAGT 1529
 Db 547 TCAAGGGGCTTGGCTTCAAACTGCTTTTCCAGGCTATCTACTCAGAGAAAGATAAAAGT 488
 QY 1530 GTGATCTAAGAAAAGATGCTTTTAGGAAAGTGAATAATTTTGTGTTTGTATTTGA 1589
 Db 487 GTGATCTAAGAAAAGTATGCTTTTAGGAAAGTGAATAATTTTGTGTTTGTATTTGA 428
 QY 1590 AGAAGAATGATGCTTTTGAACAAGAAATCATATATGATGATATATTTTAAAGTATT 1649
 Db 427 AGAAGAATGATGCTTTTGAACAAGAAATCATATATGATGATATATTTTAAAGTATT 368
 QY 1650 TGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAATAATGCTTTGGT 1709
 Db 367 TGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAATAATGCTTTGGT 308
 QY 1710 TTTCAATTTGCTTACCAAAAAACAACAAGAAAGTCTCTTTTGAAGACTTCCACCTG 1769
 Db 307 TTTCAATTTGCTTACCAAAAAACAACAAGAAAGTCTCTTTTGAAGACTTCCACCTG 248
 QY 1770 CTCCTATGCTGGTACCTGAGTCAAAATTTGTCATTTTGTCTGTGAAAAATAAATTTCT 1829
 Db 247 CTCCTATGCTGGTACCTGAGTCAAAATTTGTCATTTTGTCTGTGAAAAATAAATTTCT 188
 QY 1830 TCTGTACCATTTCTGTTTGTAGTTTACTTAAATCTGTAATCTGTAATCTGTTTCTGTTAT 1889
 Db 187 TCTGTACCATTTCTGTTTGTAGTTTACTTAAATCTGTAATCTGTAATCTGTTTCTGTTAT 128
 QY 1890 TCCAAATTTGATGAATCGAATCCAAATTTGAAAGTTTGTGCGACGCTGCTGCTAGCTT 1949
 Db 127 TCCAAATTTGATGAATCGAATCCAAATTTGAAAGTTTGTGCGACGCTGCTGCTAGCTT 68
 QY 1950 AAATGAATGCTTCTATTTGCTTTATACATTTATATTAATTAATTTTCTAT 2009
 Db 67 AAATGAATGCTTCTATTTGCTTTATACATTTATATTAATTAATTTTCTAT 8
 QY 2010 T 2010
 Db 7 T 7

RESULT 10
 AUI47533/c
 LOCUS
 DEFINITION AUI47533 MAMMA1 Homo sapiens CDNA clone MAMMA1000956 3', mRNA
 sequence.
 ACCESSION AUI47533
 VERSION AUI47533.1 GI:11009054
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 577)
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
 Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T.,
 Sugano,S. and Isogai,T.).
 HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
 Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M.,
 Nakamura,Y., Nagai,T., Sugano,S., Isogai,T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
 Research Institute; cDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES
 source
 Location/Qualifiers
 1. 577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MAMMA1000956"
 /tissue_type="mammary gland"
 /clone_lib="MAMMA1"
 /notes="Vector: pME18SFL3"

ORIGIN
 Query Match 26.7%; Score 535.8; DB 9; Length 577;
 Best Local Similarity 97.2%; Pred. No. 5.3e-97;
 Matches 562; Conservative 0; Mismatches 14; Indels 2; Gaps 2;
 QY 1265 AAAACAGCTTAGGGATTAAATGCTCCTCATTTATAATGAAGATTAATAATGAAGCTTTAAT 1324
 Db 577 AAAACAGCTTAGGGATTAAATGCTCCTCATTTATAATGAAGATTAATAATGAAGCTTTAAT 518
 QY 1325 CAGCATTTGAAGAAATTAATGCTTCTCATATGCTGTTTTCAGCTAGGAGTTAG 1384
 Db 517 CAGCATTTGAAGAAATTAATGCTTCTCATATGCTGTTTTCAGCTAGGAGTTAG 459
 QY 1385 AAATCCTAACTTCTTATCTCTCTCCAGAGGC-TTTTTTTTCTTCTGTTATTAAT 1443
 Db 458 AAATCCTAACTTCTTATCTCTCTCCAGAGGC-TTTTTTTTCTTCTGTTATTAAT 399
 QY 1444 AACATTTTAAACCGAGATATTTTGTCAAGGGCTTTGCAATCAAACTGCTTTTCCAGG 1503
 Db 398 AACATTTTAAACCGAGATATTTTGTCAAGGGCTTTGCAATCAAACTGCTTTTCCAGG 339
 QY 1504 GCTATCTCAGAGAAAGATATAAGTGTGATCTAAGAAAAGTGTGTTTGTAGGAAAGT 1563
 Db 338 GCTATCTCAGAGAAAGATATAAGTGTGATCTAAGAAAAGTGTGTTTGTAGGAAAGT 279
 QY 1564 GAAAAATATTTTCTTTTGTATTTGAAGAAAGATGATGATTTTCAAGAAAATCATATA 1623
 Db 278 GAAAAATATTTTCTTTTGTATTTGAAGAAAGATGATGATTTTCAAGAAAATCATATA 219
 QY 1624 TGATGATATATTTTAAATAGTATTGAGTACAGACTTTGAGGTTTCAATCAATATAAT 1683
 Db 218 TGATGATATATTTTAAATAGTATTGAGTACAGACTTTGAGGTTTCAATCAATATAAT 159
 QY 1684 AAAAGAGCAGAAAAATATGCTTGTGTTTTCATTTGCTTACCAAAAAACAACAAAAA 1743
 Db 158 AAAAGAGCAGAAAAATATGCTTGTGTTTTCATTTGCTTACCAAAAAACAACAAAAA 99
 QY 1744 AGTTGCTTTGAGAACTTCACTGCTCTCTATGTTGGTACCTGAGTCAAAATTTGTCATT 1803
 Db 98 AGTTGCTTTGAGAACTTCACTGCTCTCTATGTTGGTACCTGAGTCAAAATTTGTCATT 39
 QY 1804 TTTGTTCTGTGAAAAATAAATTTCTTCTTGTACCAATT 1841
 Db 38 TTTGTTCTGTGAAAAATAAATTTCTTCTTGTACCAATT 1

RESULT 11
 CB851224
 LOCUS
 DEFINITION CB851224 adn-j-21-0-UI-s1 UI-CF-DUI Homo sapiens CDNA clone
 UI-CF-DUI-adn-j-21-0-UI 3', mRNA sequence.
 ACCESSION CB851224
 VERSION CB851224.1 GI:30045993
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 551)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction; two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL

[illegible]

```
Db 576 TTTTGGTTGCAACGAAAGCGAGTAGCTACAGATACCTCGGATACCTCCCTTCGCACA 517
QY 709 ACCCAAAAAAGTTATCACACCGGAAAGAGTACCGAGCGCTCTACTCCAGAACTCAGTAT 768
Db 516 ACCCAAAAAAGTTATCACACCGGAAAGAGTACCGAGCGCTCTACTCCAGAACTCAGTAT 457
QY 769 GGTAGTGTGTATG-TTTTTTAACTTTACTATATAAGCCATGCAAAATGCAAAAATCTA 827
Db 456 GGTAGTGTGTATG-TTTTTTAACTTTACTATATAAGCCATGCAAAATGCAAAAATCTA 397
QY 828 TATTACTTCTCAAAATGCAACCCCAAGAGAACTTTGATTTACTGTTCTTAACGTCCTAAT 887
Db 396 TATTACTTCTCAAAATGCAACCCCAAGAGAACTTTGATTTACTGTTCTTAACGTCCTAAT 337
QY 888 CTTAATTACAGAACTGTCATCAGCTATTATGATTCATTAAGCTATTTCAGCAAGATG 947
Db 336 CTTAATTACAGAACTGTCATCAGCTATTATGATTCATTAAGCTATTTCAGCAAGATG 277
QY 948 AGATATTAAACCAATGCTTTGATTTCTAGAAAGTATAGTAATTTGTTTCTAAGGNG 1007
Db 276 AGATATTAAATCCAAATGCTTTGATTTCTAGAAAGTATAGTAATTTGTTTCTAAGGNG 217
QY 1008 GTTCAAGCATCTCTTTTTTATCAATTTACTTCAAAATGCAATTTGCTTAAAGCTGCATTA 1067
Db 216 GTTCAAGCATCTCTTTTTTATCAATTTACTTCAAAATGCAATTTGCTTAAAGCTGCATTA 157
QY 1068 TTTTACTACTGTAATTTCTCCAGACATAGCAATTAATGATAGATGAGTGTAACATTTA 1127
Db 156 TTTTACTACTGTAATTTCTCCAGACATAGCAATTAATGATAGATGAGTGTAACATTTA 97
QY 1128 TATCTCACATAGACATGCTTATATGTTTATTTAAATGAAATGCCAGTCCATTA 1187
Db 96 TATCTCACATAGACATGCTTATATGTTTATTTAAATGAAATGCCAGTCCATTA 37
QY 1188 CTGAATAA 1195
Db 36 CTGATAAA 29
```

```
RESULT 13
AW237740/c
LOCUS xmb1a03.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2690572 3',
DEFINITION mRNA sequence.
ACCESSION AW237740
VERSION AW237740.1 GI:6570129
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 513)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 462.
FEATURES
Location/Qualifiers
1..513
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

```
/clone="IMAGE:2690572"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid11"
/notes="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtracive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
```

ORIGIN

```
Query Match 25.5%; Score 512; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. NO. 3.2e-92;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1499 CCAGGGCTATCTCAGAGAAAGATAAAGTGTGATCTAAGAAAGTGAAGTGGTTTAGG 1558
Db 513 CCAGGGCTATCTCAGAGAAAGATAAAGTGTGATCTAAGAAAGTGAAGTGGTTTAGG 454
QY 1559 AAAGTGAATAATATTTTGTGTTTGTATTTGAAGAAAGATGATGCAATTTTGACAAGAAATC 1618
Db 453 AAAGTGAATAATATTTTGTGTTTGTATTTGAAGAAAGATGATGCAATTTTGACAAGAAATC 394
QY 1619 ATATATGTATGGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGGTTTCATCAATA 1678
Db 393 ATATATGTATGGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGGTTTCATCAATA 334
QY 1679 TAAATAAAGAGCAGAAATAATATGCTTGGTTTTCATTTGCTTACCAAAAAACAAAC 1738
Db 333 TAAATAAAGAGCAGAAATAATATGCTTGGTTTTCATTTGCTTACCAAAAAACAAAC 274
QY 1739 AAAAAAGTTGCTCTTTGAGAACTTCACCTGCTCTATATGTGGGTACCTGAGTCAAAATG 1798
Db 273 AAAAAAGTTGCTCTTTGAGAACTTCACCTGCTCTATATGTGGGTACCTGAGTCAAAATG 214
QY 1799 TCATTTTGTGTCGAAAAATAAATTTCCCTTCTGTACCAATTTCTGTTAGTTTACTA 1858
Db 213 TCATTTTGTGTCGAAAAATAAATTTCCCTTCTGTACCAATTTCTGTTAGTTTACTA 154
QY 1859 AAATCTGTAATAACTGATATTTTCTGTTTATTTCCAAATTTGATGAACCTGCAATCCAAAT 1918
Db 153 AAATCTGTAATAACTGATATTTTCTGTTTATTTCCAAATTTGATGAACCTGCAATCCAAAT 94
QY 1919 TTGAAAGTTTGTGTCGACGCTGCTAGCTTAAATGAATGTTCTTATTTGCTTTATACA 1978
Db 93 TTGAAAGTTTGTGTCGACGCTGCTAGCTTAAATGAATGTTCTTATTTGCTTTATACA 34
QY 1979 TTTATATTAATAAATTTGTACATTTTCTAAT 2010
Db 33 TTTATATTAATAAATTTGTACATTTTCTAAT 2
```

RESULT 14

```
BF675808
LOCUS 602083725F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4247718 5',
DEFINITION mRNA sequence.
ACCESSION BF675808
VERSION BF675808.1 GI:11949703
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 689)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI065 row: b column: 07
High quality sequence stop: 676.
Location/Qualifiers

source
1. .689
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4247718"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: Sfii (ggcgcctcgcc); Site 2: Sfii (ggcctatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 25.1%; Score 505.2; DB 10; Length 689;
Best Local Similarity 95.8%; Pred. No. 6.9e-91;
Matches 638; Conservative 0; Mismatches 13; Indels 15; Gaps 11;
QY 943 GAATGAGATTAAACCCATGCTTTGATGTTCTAGAAAGTAGTAATTTGTTTCTA 1002
DB 1 GAATGAGATTAAACCCATGCTTTGATGTTCTAGAAAGTAGTAATTTGTTTCTA 57
QY 1003 AGGTGGTTCAGCATCTCTTTTATCAATTTCTCAAAATGACATGCTAAAGACTG 1062
DB 58 AGGTGGTTCAGCATCTCTTTTATCAATTTCTCAAAATGACATGCTAAAGACTG 116
QY 1063 CATTATTACTAGTAAATTTCTCAGACATAGCATATGTCATAGATGAGTGAAC 1122
DB 117 CATTATTACTAGTAAATTTCTCAGACATAGCATATGTCATAGATGAGTGAAC 176
QY 1123 ATTATATCTCAGATGACATGTTATATGTTTATTTTAAATGAATGCCAGTCCA 1182
DB 177 ATTATATCTCAGATGACATGTTATATGTTTATTTTAAATGAATGCCAGTCCA 236
QY 1183 TTACTGTAATAATAGAACTCAACTATTGCTTTTCAGGAAATCATGATAGGTTCAA 1242
DB 237 TTACTGTAATAATAGAACTCAACTATTGCTTTTCAGGAAATCATGATAGGTTCAA 296
QY 1243 GAAGTTACTATTAATGTTTAAACACAGCTTAGGGATTAATGCTCCATTAATGA 1302
DB 297 GAAGTTACTATTAATGTTTAAACACAGCTTAGGGATTAATGCTCCATTAATGA 356
QY 1303 AGATTAAATGAAGGCTTTAAATCAGCATTTAAAGAAATGAATGGCTTTCTCATATGC 1362
DB 357 AGATTAAATGAAGGCTTTAAATCAGCA-TGTAAGGAATTAATGAGG-TTTCGTATGC 414
QY 1363 TGTTTTTACCTAGAGTTAGAAATCCTAACTCTTTATCTCTCTCCAGAGGCTTT 1422
DB 415 TGTTTTTACCTAGAGTTAGAAATCCTAA-TTCTTTATCTCTCTCCAGAGGCTTT 473
QY 1423 TTTTCT-TGTGTATTAAATTAACATTTT-AAAACGACAGATTTTGTCAAGGGGCTT 1480
DB 474 TTTTCTCTGTGTATTAAATTAACATTTTAAACAGCAGATTTTGTCAAGGGGCTT 533
QY 1481 TGCATTCAACATGCTTTTCAGGGCTTACTACTCAGAGAAAGATAA-AAAGTGTGATCTAAG 1539

DB 534 GTGATTCAAACTGCTTTTCCAGGGCTACTACTCAGAGAAAGATAACAGTGTGATCTAAG 593
QY 1540 AAAAAGTGTGATGTTTGTAGAAAGTGAATAATTTTGTGTTTGTATTGAAGAAGATGA 1599
DB 594 AAAAAGTGTGATG-TTTAGAAAGTGAATAATTTTGTGTTTGTATTGAAGAAGATGA 648
QY 1600 TGCATT 1605
DB 649 TGTCTT 654
AY400667 678 bp DNA linear GSS 15-DEC-2003
Mus musculus CLDN8 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY400667
VERSION AY400667.1 GI:39756656
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 678)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Gargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 678)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Gargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .678
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1. ->678
/gene="CLDN8"
/locus_tag="HCM0633"
ORIGIN
Query Match 24.6%; Score 494; DB 29; Length 678;
Best Local Similarity 83.0%; Pred. No. 1.2e-88;
Matches 563; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 97 ATGGCAACCCATGCCCTTGAATCGTGGGCTGTTCTTGGTGGTGGTGGTGGTGGTGGG 156
DB 1 ATGGCAACCCATGCCCTTGAATCGTGGGCTGTTCTTGGTGGTGGTGGTGGTGGTGGG 60
QY 157 ACAGTGGCTGTCACTGTGTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAACCAATC 216
DB 61 ACAGTGGCTGTGACTATCATGCCTCAGTGGAGAGTGTGGCTTCATCGAAGTAACATT 120
QY 217 GTGGTTTGTGAAATCTTCGGAGGAGTGTGATGATTCGTCGAGGAGGCTTAACATC 276
DB 121 GTGGTTTGTGAAACCGCTGGGAGGCTTGTGGATGAATGATGAGGCAATGCCAATC 180
QY 277 AGGATGAGTGAATAATCTATGATTCCTGCTGCTCTTTCTCCGACCTACAGGAGCC 336
DB 181 AGAATGAGTGAAGGTCTACGACTCCCTGCTGCTCTTAGTCCAGACCTCCAGGATCC 240
QY 337 AGAGGAGTGAATGTGCTGCTTCCGTGATGCTCTTCTTGGCTTTTCATGATGAGCCATCCTT 396

241	Db	CGAGCAGTGAATGTGTGCTCGCGTCTTGCGCTTTCTTGGCTTTTCATGACAGCCATCCCTC	300
397	Qy	GGCATGAATGCACCAAGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTTCTGTGTG	456
301	Db	GGATGAAGTGCACCAAGTGCACGGGGGACGATGAGAACGTGAAGAGCGCATCTTCTGTGTG	360
457	Qy	AGGGTGGAAATCATCTTCATCATCACGGGCATGGTGTGCTCATCCCTGTGAGCTGGGTT	516
361	Db	ACAGCGGAATCATCTTCTTCATCACGGCTTGGTGTGCTCATCCCTGTACGTGGGTT	420
517	Qy	GCCAATGCCATCATCAGAGATTCTATACTCAATAGTGAATGTTGCCCAAAAACGTGAG	576
421	Db	GCCAAATCCATCATCAGAGACTTCTACAACCCACTGGTGGATGTGGCCCTTAAGCGCGAG	480
577	Qy	CTTGGAGAAGCTCTCTACTTAGTAGTGACACGGCACTGGTGTGCTGAATGTTGGAGGAGCT	636
481	Db	CTGGGAGAAGCCCTCTACATAGGTGGACACACGCGCTGGTGTGATCTCGTGGAGGACGA	540
637	Qy	CTGTTCTGCTGCGTTTTTTTGTTCGAAACGAAAGAGCAGTAGCTACAGATACTCGATACT	696
541	Db	CTGTTCTGTTGTGTGTTTTTGTGTTACTGTAAGGAGCAACAGTTTACAGGTACTTCGGTACCA	600
697	Qy	TCCCATCGCACCAACCCAAAAAGTTATCACACCGGAAGAAGTCAACCGAGCGCTTACTCC	756
601	Db	TCCCATCGCAACCACTCACGGAGTTTCCACGCCGAAAAGAGATCTCCGAGCATATACTCC	660
757	Qy	AGAAGTCAGTATGTGTAG	774
661	Db	AAAAGTCAGTATGTGTAG	678

RESULT 16	AW205135/c
LOCUS	AW205135/c
DEFINITION	UI-H-B11-aem-g-01-0-UI.s1 NCI CGAP_Sub3 Homo sapiens cDNA clone IMAGE:720040 3', mRNA sequence.
ACCESSION	AW205135
VERSION	AW205135.1
KEYWORDS	GI:504607
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap , National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image.html Seq primer: M13 Forward POLYA=yes.

ORIGIN

Query Match	23.9%	Score	480.4	DB	10	Length	492
Best Local Similarity	99.8%	Pred. No.	6.9e-86				
Matches	481	Conservative	0	Mismatches	1	Indels	0
Gaps	0						
QY	1219	ACGGAAATCATGGATAGGGTTGAAGAAGCTTACTATTGTTTAAAAACAGCTTAGGG	1278				
DB	492	ACGGAAATCATGGATAGGGTTGAAGAAGGTTACTATTGTTTAAAAACAGCTTAGGG	433				
QY	1279	ATTATGTCCTCCATTTTATAATCAAGATTAAAAATGAAGGCTTTAATCAGCATTTGTAAGG	1338				
DB	432	ATTATGTCCTCCATTTTATAATCAAGATTAAAAATGAAGGCTTTAATCAGCATTTGTAAGG	373				
QY	1339	AAATGAATGCTCTTCGATATGCTGTTTTTAGCTAGGAGTTAGAAAATCCTAATTCT	1398				
DB	372	AAATTGAATGCTCTTCGATATGCTGTTTTTAGCTAGGAGTTAGAAAATCCTAATTCT	313				
QY	1399	TTATCCTCTCTCCAGAGGCTTTTTTTTCTTGTTGTTAAATTTAAACATTTTTTAAACG	1458				
DB	312	TTATCCTCTCTCCAGAGGCTTTTTTTTCTTGTTGTTAAATTTAAACATTTTTTAAACG	253				
QY	1459	CAGATATTTTGTCAAGGGGCTTTTGCATTCAAACTGCTTTTCCAGGGCTATATCTCAGAAGA	1518				
DB	252	CAGATATTTTGTCAAGGGGCTTTTGCATTCAAACTGCTTTTCCAGGGCTATATCTCAGAAGA	193				
QY	1519	AAGATAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTGAATAATTTTGT	1578				
DB	192	AAGATAAAGTGTGATCTAAGAAAAAGTGTGGTTTTAGGAAAGTGAATAATTTTGT	133				
QY	1579	TTTGTATTTTGAAGAAGAAATCATGCAATTTTGAACAAGAAATCATATATGTATGGATATATTT	1638				
DB	132	TTTGTATTTTGAAGAAGAAATCATGCAATTTTGAACAAGAAATCATATATGTATGGATATATTT	73				
QY	1639	TAATAAGTATTTTGAGTACAGACTTTTCAGGTTTTCATCAATATAAAATATAAAGCAGCAAAA	1698				
DB	72	TAATAAGTATTTTGAGTACAGACTTTTCAGGTTTTCATCAATATAAAATATAAAGCAGCAAAA	13				
QY	1699	TA 1700					
DB	12	TA 11					

```
RESULT 17
CB418866
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 683)
Smith, T.P.L., Roberts, A.J., Echternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keele, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: FQY8011 row: A column: 23
Seq primer: GTAATACGACTCACTATAGG.
Location/Qualifiers
1. .681
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARCB 6BOV"
/notes="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

FEATURES
source
294 CTATGATTCCTGCTGCTCTTTCTCGGACCTACAGGACCCAGAGACTGATGTGTC 353
1 CTACGACTCGCTGCTGCTCTCTCTCGGACCTACAGGACCCAGAGACTGATGTGTC 60
354 TGCTTCGGTGATGCTCTCTGCTTTCATGATGCCATCTTGGCATGAATGCACAG 413
61 CGCTCGGTGCTGCGCTCTTGGCTTCTCGGCGCTCTCGGCGATGAAGTGATCCAG 120
414 GTGCACGGGGGCAATGAGAAGTGAAGGCTCACATTTCTGCTGACGGCTGGAATCATTT 473
121 ATGCGCGGGGACGACGACAAAGTGAAGGTCACATTTCTGCTGACCGCTGGAGTATTT 180
474 CATCATCAGGGCATGCTGGTGTCTATCCCTGTAGCTGGGTGCGCATGCGCATCATCAG 533
181 CATCATCTGCGCTCGTGGTGTCTATCCCGCTGAGCTGGGTGCGCATTCATCATCAG 240
534 AGATTTCTTAATCAATAGTGAATGTTGCCCAAAAGCTGAGCTTGGAGAAGCTCTCTA 593
241 AGACTTCTACACCCCAATAGTGGATATTGCCAGAAACGGGAGCTGGAGAGCCCTCTA 300
594 CTTAGATGAGACACCGCATGCTGGTGTATTTGTGAGAGAGCTGTGTTCTGCTGGTTTT 653
301 CATAGGCTGAGACACCGCCCTGGTGTGATTTGTGAGGGGGGCTGTCTGTGCGTTTC 360
654 TTGTTCCAGAAAGAGAGAGTAGCTACAGATACCTCCATCTCCCATCCGACACCCA 713

ORIGIN
Query Match 23.3%; Score 468.4; DB 14; Length 683;
Best Local Similarity 83.1%; Pred. No. 1.6e-83;
Matches 570; Conservative 0; Mismatches 111; Indels 5; Gaps 3;

QY 294 CTATGATTCCTGCTGCTCTTTCTCGGACCTACAGGACCCAGAGACTGATGTGTC 353
DB 1 CTACGACTCGCTGCTGCTCTCTCTCGGACCTACAGGACCCAGAGACTGATGTGTC 60
QY 354 TGCTTCGGTGATGCTCTCTGCTTTCATGATGCCATCTTGGCATGAATGCACAG 413
DB 61 CGCTCGGTGCTGCGCTCTTGGCTTCTCGGCGCTCTCGGCGATGAAGTGATCCAG 120
QY 414 GTGCACGGGGGCAATGAGAAGTGAAGGCTCACATTTCTGCTGACGGCTGGAATCATTT 473
DB 121 ATGCGCGGGGACGACGACAAAGTGAAGGTCACATTTCTGCTGACCGCTGGAGTATTT 180
QY 474 CATCATCAGGGCATGCTGGTGTCTATCCCTGTAGCTGGGTGCGCATGCGCATCATCAG 533
DB 181 CATCATCTGCGCTCGTGGTGTCTATCCCGCTGAGCTGGGTGCGCATTCATCATCAG 240
QY 534 AGATTTCTTAATCAATAGTGAATGTTGCCCAAAAGCTGAGCTTGGAGAAGCTCTCTA 593
DB 241 AGACTTCTACACCCCAATAGTGGATATTGCCAGAAACGGGAGCTGGAGAGCCCTCTA 300
QY 594 CTTAGATGAGACACCGCATGCTGGTGTATTTGTGAGAGAGCTGTGTTCTGCTGGTTTT 653
DB 301 CATAGGCTGAGACACCGCCCTGGTGTGATTTGTGAGGGGGGCTGTCTGTGCGTTTC 360
QY 654 TTGTTCCAGAAAGAGAGTAGCTACAGATACCTCCATCTCCCATCCGACACCCA 713

RESULT 18
CF766535
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 681)
Wang, Y.H., McWilliam, S. and Lehnert, S.
Transcription profiling of cattle skin
Unpublished (2003)
Contact: Dr Yonghong Wang
Functional Genomics Lab
CSIRO Livestock Industries
Level 5, Queensland Biosciences Precinct, University of Queensland,
306 Carmody Road St. Lucia QLD Australia
Tel: 07 3214 2445
Fax: 07 3214 2685
Email: Yonghong.Wang@csiro.au
Plate: 43 row: B column: 11.
Location/Qualifiers
1. .681
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone_lib="CCLO03314"
/sex="female"
/tissue_type="pooled"
/dev_stage="Adult"
/lab_host="XLI-BlueMRF, strain"
/clone_lib="Bos taurus skin cDNA library"
/notes="Organ: skin; Vector: Uni-ZAPXR; Site 1: EcoRI;
Site 2: Xho I; Library made from pooled skin of adult
female Hereford-Shorthorn."

FEATURES
source
1 GGAAAACTGTCTCTCTTGTGGCAGAGAACCTGTTCAGAGAGCCTTACCAGTC 60
13 GGAAAACTGTCTCTCTCGGGCGCAGAGAACCTGTTCAGAGAGCCTTACCAGTC 72

ORIGIN
Query Match 23.0%; Score 462.2; DB 14; Length 681;
Best Local Similarity 80.6%; Pred. No. 2.8e-82;
Matches 539; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GGAAAACTGTCTCTCTTGTGGCAGAGAACCTGTTCAGAGAGCCTTACCAGTC 60
DB 13 GGAAAACTGTCTCTCTCGGGCGCAGAGAACCTGTTCAGAGAGCCTTACCAGTC 72
```

QY 61 CGGAGTCAGCTGGCTAAATCTATCCAGAGGATATGCAACCCATCGCTAGAAATC 120
 Db 73 AGGCGCCCTGGTCCGATAAACCCAACTGGAGGAGATGGCTACCTACGCCCTGCAATC 132
 QY 121 GCTGGGCTGTTTCTTGTGGTGTGGTAAATGGTGGGACAGTGGCTGTCTCATGTGCTGCCT 180
 Db 133 GCCGACCTGGTCTCGTGGTGTGGGATGGAGGACACAGTGGCTGCCAGGCTCATGCTC 192
 QY 181 CAGTGGAGAGTGTGGGCTTCTATTTGAAACCAACATCGTGGTTTTTGAACACTTCTGGAA 240
 Db 193 CAGTGGAGAGTGTGGGCTTCTATTTGAAAGCAATATGGGGTCTTTGAAACCTCTGGAA 252
 QY 241 GGACTGTGGATGAATGGTGGAGGAGCTAAATCAGGATGAGTGCAGGAAATCTATGAT 300
 Db 253 GGACTATGATGAGTTCATGAGGAGCTATCAATCACAATGAGTGCAATCTACGAC 312
 QY 301 TCCCTGTGGCTTCTTCTCCGACCTACAGGAGCCAGAGGAGTGTGCTGCTTCC 360
 Db 313 TCGCTGTGGCTCTCTCTCCGACCTACAGGAGCCAGAGGAGTGTGCTGCGGCTCG 372
 QY 361 GTGATGCTCTTCTGGCTTCTATGATGCTGCTTCCGATGAGTGAATGCAAGGTCAGC 420
 Db 373 GTGCTGGCTGTGGCTTCTTCTGACGGCCGACTCGGCAATGAGTGTACGAGTCCGCC 432
 QY 421 GGGGCAATGAGAGGAGGCTCAATCTCTGCTGACGGCTGGAATCATCTTCATCATC 480
 Db 433 GGAGACGACGACAAAGGTGAAAGGTCAATCTGCTGACCGCTGGAGTGAATCATCATC 492
 QY 481 ACGGCATGCTGTGCTCATCCCTGTGAGTGGTGGTGGCCATGCCATCATCAGAGATTTC 540
 Db 493 ACTGCGCTCGGGTCTCATCCCGAGAGCTGGTGGTGGGTTGCCAATTCATCATCAGAGACTTC 552
 QY 541 TATACTCAATAGTGAATGTTGCCCAAAACGCTGAGCTTGGAGAAGCTCTCTACTTAGGA 600
 Db 553 TACAACCAATAGTGGATATGTTCCAGAAACGGGAGCTGTGAGAAGCCCTCTACATATGC 612
 QY 601 TGGACACGCGACCTGGTCTGATTTGGAGGAGCTCTGTCTGCTGGTTTTTTTGTTC 660
 Db 613 TAGACCATGACCTGGTGTGCTGATGAGGAGGCGCTGTCTGTGGCTTCCCTTTC 672
 QY 661 AACGAAAG 669
 Db 673 CATGAAAG 681

RESULT 19
 CFI69716
 LOCUS
 DEFINITION
 B0817B10-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
 musculus cDNA clone NIA:B0817B10 IMAGE:30469077 5', mRNA sequence.
 ACCESSION
 CFI69716
 VERSION
 CFI69716.1 GI:33279265
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 633)
 Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 JOURNAL
 MEDLINE
 PUBMED
 21429098
 11544199
 COMMENT
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0817 row: B column: 10
 Seq primer: M13 Reverse

High quality sequence stop: 633
 POLYA=No.
 Location/Qualifiers
 1..633
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:B0817B10-5"
 /db_xref="taxon:10090"
 /clone="NIA:B0817B10 IMAGE:30469077"
 /dev_stage="Newborn Kidney"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long 1)"
 /notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). In brief, double-stranded cDNAs were synthesized with an oligo(dT) primer (Invitrogen): 5'-TGACTAGTTCATGATCGAGCGGCGCCCTTTT-3' from 26 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 22.9%; Score 460.2; DB 14; Length 633;
 Best Local Similarity 82.9%; Pred. No. 7.2e-82;
 Matches 525; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 98 TGGCAACCCATGCTTAGAAATCGTGGGCTGTTCCTGGTGGTGTGGAATGTTGGGCA 157
 Db 1 TGGCAACCTACGCTCTCAATGGCTGCACCTGGTGGTGGTGGTGGTGGGCA 60
 QY 158 CAGTGGCTGTCACTGTCACTGCCTCAGTGAGAGTGTGGCCCTTCATTGAAACACATCG 217
 Db 61 CGGTGGCTGTGACTATCATGCCCTCAGTGAGAGTGTCTGCCCTTCATCGAATACATG 120
 QY 218 TGGTTTTTGAACACTTCTGGGAGGACTGTGGATGAATTCGTGAGGAGGCTAACATCA 277
 Db 121 TGGTGTGAGAACCCCTGGGAGGCTTGTGATGAATGTATGAGGATGCCAATCA 180
 QY 278 GGATGAGTCAAAATCTATGATTCCTCTGGCTCTTCTCCGAGCTACAGGAGCA 337
 Db 181 GAATGAGTCAAGGCTACGACTCCCTCTGGCTTTAGTCCAGACCTCCAGGATCCC 240
 QY 338 GAGGACTGATGTGTGCTTCCGTGATGCTCTTCTGGCTTTTCATGATGGCATCCTTG 397
 Db 241 GAGGACTGATGTGTGCTGGCTCGTCTTCTTGGCTTTTCATGAGCAATCCTCG 300
 QY 398 GCATGAATGCACAGGTGACGGGGACAATGAGAGGTGAAGGCTCATCTTCTGCTGA 457
 Db 301 GAATGAAGTCCACAGATGCACGGGGAGCATGAGAACGTCGATCTTGTCTGA 360
 QY 458 CGGCTGAATCATCTTTCATCATCCGGGATGTGTGCTCATCTCCCTGTGAGCTGGTTG 517
 Db 361 CAGCCGGAATCATCTTCTTTCATCACCGGTTTGGTGTGCTCATCTCCTGTAGTGGTTG 420
 QY 518 CCAATGCCATCATCAGAGATTTCTAATCAATAGTGAATTTGCCCAAAACCTGAGC 577
 Db 421 CCAATTCATCATCAGAGACTTCTCAACCCCACTGGTGGATGTGGCCCTTAAGCCGAGC 480
 QY 578 TTGGAGAGGCTCTCTACTTAGGATGGACCAACGCACTGGTGTGCTGATTGTTGGAGGCTC 637


```
Db      481 TGGAGAGCCCTCTACATAGGCTGGACACACAGCGCTGGTCTGATCGCTGGAGGAC 540
QY      638 TGTTCTGCTGGTTTTTTTTCGACAGAAAGAGAGAGTAGCTACGATACCTGATACCTT 697
Db      541 TGTCTGTTGTGTGTTTCTTCTGTAAGAGAGCAACAGTTACAGGTACTCGTACCAT 600
QY      698 CCATCGCACACCCCAAAAAGTTATCACCG 730
Db      601 CCCATCGCACCACTACACGGAGTTCCAGCGCG 633

RESULT 20
BG402412
LOCUS   BG402412.1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594155 5',
DEFINITION mRNA sequence.
ACCESSION BG402412
VERSION   BG402412.1 GI:132995860
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 809)
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: CLONTECH Laboratories, Inc.
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLM1335 row: e column: 04
          High quality sequence stop: 567.
          Location/Qualifiers
            1..809
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:4594155"
              /lab_host="NIH MGC_75"
              /clone_lib="NIH MGC_75"
              /note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site 1:
              SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctcgcc); 5' and
              3' adaptors were used in cloning as follows: 5' adaptor
              sequence: 5'-CACGCCATATGACC-3' and 3' adaptor sequence:
              5'-ATTCTAGAGCCGAGCGCCGACARG-dt(30)BN-3' (where B = A,
              C, or G and N = A, C, G, or T). Average insert size 1.65
              kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
              by PCR. This library was enriched for full-length clones
              and was constructed by Clontech Laboratories (Palo Alto,
              CA). Note: this is a NIH_MGC Library."

FEATURES
source
  1..809
    Query Match 22.1%; Score 445.2; DB 12; Length 809;
    Best Local Similarity 96.1%; Pred. No. 6.9e-79;
    Matches 490; Conservative 0; Mismatches 13; Indels 7; Gaps 3;

QY      1 GGAAACATCTTCTCTCTGTCGACAGAGACCTGCTTCAACGACAGTACGATTC 60
Db      15 GGAAACATCTTCTCTCTGTCGACAGAGACCTGCTTCAACGACAGTACGATTC 74
QY      61 CGAGTCCAGCTGGCTAAACATCATCCAGAGGATAATGGCAACCCATGCCCTTGAATC 120
Db      75 CGAGTCCAGCTGGCTAAACATCATCCAGAGGATAATGGCAACCCATGCCCTTGAATC 134
QY      121 GCTGGCTGTTTCTTGGTGGTGGATGGTGGGACAGTGGCTGCTGATCGCT 180
```

```
Db      135 GCTGGGCTGTTTCTTGGTGGTGTGGAATGCTGGGACAGTGGCTGCTGCTCATGCT 194
QY      181 CAGTGGAGAGTGTGGGCTTCATTGAAACAAACATCGTGGTTTTTGAACACTTCTGGGAA 240
Db      195 CAGTGGAGAGTGTGGGCTTCATTGAAACAAACATCGTGGTTTTTGAACACTTCTGGGAA 254
QY      241 GGACTGTGATGAATGGTGTGAGGAGGCTAACATCAGATGCAATGCAAAATCTATGAT 300
Db      255 GGACTGTGATGAATGGTGTGAGGAGGCTAACATCAGATGCAATGCAAAATCTATGAT 314
QY      301 TCCCTGCTGGCTCTTTCTCCGACCTACAGCAGCAGAGGACTGATGTGCTGCTTCC 360
Db      315 TCCCTGCTGGCTCTTTCTCCGACCTACAGCAGCAGAGGACTGATGTGCTGCTTCC 374
QY      361 GTGATGTCTCTTCTGGCTTTCA-TGATGGCCATCTTGG----CATGAAATGCAACCAAGT 415
Db      375 GTGATGTCTCTTCTGGCTTTCA-TGATGGCCATCTTGGCTGTGAACATGCAACCAAGT 434
QY      416 GCACGGGGGACAATGAGAGGTGAAGGCTCACTTCTGCTGACGCTGGAATCA--TCTT 473
Db      435 GCACGGGGGACGATGAGAGGTGAAGGCTCACTTCTGCTGACGCTGGAATCACTCTTT 494
QY      474 CATCATCAGGGCATGGTGGTCTCATCCC 503
Db      495 CATCATCAGGGCATGGTGGTCTCATCCC 524

RESULT 21
BI852250
LOCUS   BI852250.1 NCI_CGAP_Man2 Mus musculus cDNA clone IMAGE:5391109 5',
DEFINITION mRNA sequence.
ACCESSION BI852250
VERSION   BI852250.1 GI:15992997
KEYWORDS EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLM11996 row: k column: 14
          High quality sequence stop: 826.
          Location/Qualifiers
            1..902
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="FVB/N-3"
              /db_xref="taxon:10090"
              /clone="IMAGE:5391109"
              /tissue_type="tumor, biopsy sample"
              /dev_stage="5 months"
              /lab_host="DH10B"
              /clone_lib="NCI_CGAP_Man2"
              /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
              Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
              library constructed by Life Technologies. Investigator
              providing samples: Gilbert Smith, NIH"

FEATURES
source
  1..902
    Query Match 22.0%; Score 443.2; DB 12; Length 902;
    Best Local Similarity 80.4%; Pred. No. 1.7e-78;

ORIGIN
```


CA313163
LOCUS 486 bp mRNA linear EST 04-NOV-2002
DEFINITION UI-CF-FNO-aex-d-08-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
CA313163
ACCESSION UI-CF-FNO-aex-d-08-0-UI 3', mRNA sequence.
VERSION CA313163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequences: 404-466, >AT-rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source Location/Qualifiers
1..486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-aex-d-08-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pVT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized Human lung epithelial cell libraries (EN1 and
DUI) The library was subtracted according to according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG ISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-FNO
TAG_SEQ=GGCTGTAGGC"

ORIGIN
Query Match 21.8%; Score 438.8; DB 14; Length 486;
Best Local Similarity 99.5%; Pred. No. 1.5e-77;
Matches 440; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1569 TATTTTGGTTTGTATTTGAAGAGATGATGCGATTGTGACAGAAATCATATATGAT 1628
Db TTTTGTATTTTGTATTTGAAGAGATGATGCGATTGTGACAGAAATCATATATGAT 61
QY 1629 GGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGGTTTCATCATATAAATAAAG 1688
Db GGATATATTTTAAATAGTATTTGAGTACAGACTTTGAGGTTTCATCATATAAATAAAG 121
QY 1689 AGCAGAAAAATATGCTTGGTTTTCATTTGCTTACCAAAAAAACAAACAAAAAAGTT 1748
Db 122 AGCAGAAAAATATGCTTGGTTTTCATTTGCTTACCAAAAAAACAAACAAAAAAGTT 181

QY 1749 GTCTTTGAGAACTTCACTGCTCTCTATGTGGGTACCTGAGTCAAAATTTGTCATTTTGT 1808
Db 182 GTCTTTGAGAACTTCACTGCTCTCTATGTGGGTACCTGAGTCAAAATTTGTCATTTTGT 241
QY 1809 TCTGTGAAATAATAATTTCCCTTCTGTGTACCACTTCTGTGTAGTTTACTTAAATCTGTAA 1868
Db 242 TCTGTGAAATAATAATTTCCCTTCTGTGTACCACTTCTGTGTAGTTTACTTAAATCTGTAA 301
QY 1869 APACTGTATTTTCTGTGTTTATTCCTCAATTTTCATCAAACTGCAATCCAAATTTGAAAGTTT 1928
Db 302 ATACTGTATTTTCTGTGTTTATTCCTCAATTTTCATCAAACTGCAATCCAAATTTGAAAGTTT 361
QY 1929 GTGTGAGCTGCTGTACCTAGCTTAATGAATGATGTTCTATTTTCTTATACATTTATATTA 1988
Db 362 GTGTGAGCTGCTGTACCTAGCTTAATGAATGATGTTCTATTTTCTTATACATTTATATTA 421
QY 1989 TAAATTTGACATTTTCTAATT 2010
Db 422 TAAATTTGACATTTTCTAATT 443

RESULT 24
CF796030/c
LOCUS 688 bp mRNA linear EST 21-OCT-2003
DEFINITION 892243 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF796030
VERSION CF796030.1 GI:37800603
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 688)
AUTHORS Smith,R.P.L., Fraking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smth@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: IMM8018 row: G column: 20
Seq primer: TAGAAGGCACAGTCGAGG.
Location/Qualifiers
1..688
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

FEATURES
source
1..688
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 21.3%; Score 429; DB 14; Length 688;
Best Local Similarity 81.4%; Pred. No. 1.3e-75;
Matches 561; Conservative 0; Mismatches 115; Indels 13; Gaps 5;

QY 732 AAGAGATCACCAGCGTCTACTCCAGAGTCAGTATGTGTAGTTGTATGTTT--TT 789
Db 688 AAGAGATCCCCGAGTGTGTCTCCAGAGTCAGTATGTATGTTGTGCTCTCTTT 629
QY 790 TAACTTTTACTATAAAGCCATGCAATGCAAAATCTTATTACTTTCTCAAAATGACC 849

```
Db 628 TGAACGCTAGAAAGCCACGCTGATGAGAAAAGGCTCTACTATTTTCTAAAAATGGAAC 569
QY 850 CCAAGAAACTTTGATTATCTGTTCTTAACTGCTCTAACTCTTAATTAACAGAACTGTGCAT 909
Db 568 CCAAGAAACACTGATTGCTGTTCTTGAAGCTGCTGACATTAATTAACAGAACTCTTCAT 509
QY 910 CAGCTATTATGATTCTATAGCTATTCTCAGCAGAAATGAGATATTAACCCCAATGCTTGG 969
Db 508 CAGCTCTTTATGATTCATAGTGAATTCAGCCGAAATGAGGATTAACACACA-----TTG 454
QY 970 ATTGTTCTAGAAAGTATAGTAAATTTGTTTCTTAAGGTGGTTCAAGCA-TCTACTCTTTTT 1028
Db 453 ATCGTTCTAGAAAGTACAGTAAATTTGTTTCTTAAAGGGTCAACATTTTCTCTTTT 394
QY 1029 ATCATTTTACTTCAAAATGACATTCCTAAAGACTGCTATTTTCTTACTGTAATTTCTCC 1088
Db 393 ATCAGTTTACTTCAAAATGACATTTGTTTGA-AGACAATTTTACAAACCGTGTGTTCTCT 335
QY 1089 ACGACATAGCATTAATGACATAGATGATGTAACATTTATATCTCAC-----ATAGAGACA 1144
Db 334 GTGACATAGCGTTAATGATATAGATGAGGTGACGTTTCTATCTCACATAAATAGAGATG 275
QY 1145 TGCATTATGTTTTTAAATGAATGCAATGCCAGTCCATTTACACTGAAATAATAGAACTC 1204
Db 274 GGCTTAGATGTTCTTATTTAAATGAATATCTGATTCATTACACTGAAATAATAGAACTC 215
QY 1205 AACTATTGCTTTTCAAGGAAATCATGATAGGTTGAAGAGGTTACTTAAATGTTTAA 1264
Db 214 AACTATTGCTTTTCAAGGAGCCGGGATGAGATTGAGAGGTTAAATTAATGTTTAA 155
QY 1265 AAAACAGCTTAGGATTAATGCTCCATTTAATGAAGATTAATAATGAAGCTTTAAAT 1324
Db 154 AAAACGGCTTAGTGAATGCACTTGATTTAATGAAGTTTAAATGAAGCTTTAAAT 95
QY 1325 CAGATTCTAAGGAATGGAATGCTGCTGATATGCTGTTTGTAGCTAGGATGATAG 1384
Db 94 CAGCAGTGAAGGAACTAATAGGCTTTCTGATATCTCTGTTTTCAGCTAGGATGATAG 35
QY 1385 AAATCCTAACTCTTTATCTCTCTCTCC 1413
Db 34 AAATCCTAACTCTTTATCTCTCTCTCC 6
```

```
RESULT 25
CFI69015 693 bp mRNA linear EST 25-JUL-2003
LOCUS B0808A07-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
DEFINITION musculus cDNA clone NIA:B0808A07 IMAGE:30468198 5', mRNA sequence.
ACCESSION CFI69015
VERSION CFI69015.1 GI:33278564
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 693)
Pisac.Y., Ko.N.T., Lim.M.K. and Ko.M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0808 row: A column: 07
Seg primer: M13 Reverse
High quality sequence stop: 693
POLYA=No. Location/Qualifiers
```

source

```
1. 693
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0808A07-5"
/db_xref="taxon:10090"
/clone="NIA:B0808A07 IMAGE:30468198"
/dev_stage="Newborn Kidney"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
1)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
In brief, double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from
26 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lona-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.0 kb. The library was
constructed by Yulan Piao."
```

ORIGIN

```
Query Match 21.3%; Score 427.8; DB 14; Length 693;
Best Local Similarity 83.4%; Pred. No. 2.2e-75;
Matches 486; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 90 GAGATAATGGCAACCCATGCCTAGAAATCTCGTGGCTGTTCTTGGTGGTGGAT 149
Db 111 GAGGACGATGGCAACCTACGCTCTTCAAATGGCTGCACCTGGTGTGGTGGTGGAT 170
QY 150 GGTGGGCACAGTGGCTGTCACTGTCTATGCTCAGTGGAGAGTGTGGCTTCATTGAAAA 209
Db 171 GGTGGGCACGTTGGCTGTGACTATCATGCTCAGTGGAGAGTGTGCTTCATCGAAG 230
QY 210 CAACATCGTGGTTTTTGAACCTTCTGGGAAGACTGTGGATGAATTCGGTGGAGCAGC 269
Db 231 TAACATTTGTTGTTTGAACCCGCTGGGAAGGCTTGTGGATGAATTTATGAGCATGC 290
QY 270 TAACATCAGGATCAGTGCAGAAATCATGATTCCTGTGGCTCTTCTCCGACCTACA 329
Db 291 CAACATCAGAAATCAGTGCAGAGGTCTACGACTCCCTGTGGCTCTTAGTCCAGACTCA 350
QY 330 GGCAAGCAGAGGACTGATGTGTGCTGCTTCCCTGATGCTCTTCTGGCTTTCATGATGGC 389
Db 351 GGCATCCGAGGACTGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
QY 390 CATCTTGGCATGAATGACAGGTCACGGGGGACATGAGAGGTTGAGGCTCAGAT 449
Db 411 CATCTCGGAATGAAGTGCACAGATGCACGGGGGACATGAGAGGTTGAGAGCCGAT 470
QY 450 TCTGCTGACGCTGGAATCATCTTCAATCATCGGGGACATGGTGGTGTGCTCATCCCTGTGAG 509
Db 471 CTGCTGACAGCCGGAATCATCTTCTTCATCCCGCTTGGTGTGCTCATCCCTGTGCTGAG 530
QY 510 CTGGTTGCCAATGCCATCATCAGATTTCTTAATCAATAGTGAATGTGCCCAAAA 569
Db 531 CTGGTTGCCAATTTCCATCATCAGAGCTTCTACAAACCCACCTGGTGGATGTGCCCTAAA 590
QY 570 ACGTGAGCTTGGAGAAAGCTCTTACTTAGGATGGACCCAGGCACTGGTGTGCTGATTTGG 629
Db 591 GCGGAGCTGGGAGAGCCCTCTACATAGCTGGACCAAGGCTGTGCTGATCGCTGG 650
```

FEATURES

```

QY 630 AGGAGCTGTCTCTGCGTGTCTTTTGTGCAAGAAAGAGC 672
    |||||
Db 651 AGGAGCACTGTCTCTGCTGTGTTTGTGTAC"GAAGAGC 693

RESULT 26
AA536178/c 490 bp mRNA linear EST 21-AUG-1997
DEFINITION mRNA sequence.
ACCESSION AA536178
VERSION AA536178.1 GI:2280431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 490)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing By: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 572 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 315.
FEATURES
Location/Qualifiers
1..490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:927725"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/clone_lib="NCI CGAP Co3"
/notes="Vector: p773B-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p773 vector. Library went through one round of
normalization."
ORIGIN
Query Match 21.1%; Score 425; DB 9; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.6e-75;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1586 TTCAACAGCAATGCATTTTGACAGCAATCATATATGATATATTTAATAAG 1645
    |||||
Db 425 TTGAAGAAATCATGCATTTTGACAGCAATCATATATGATATATTTAATAAG 366
    |||||

QY 1646 TATTGAGTACAGACTTTGAGGTTTCATCAATATAATAAAGACAGAAAAATATGCT 1705
    |||||
Db 365 TATTGAGTACAGACTTTGAGGTTTCATCAATATAATAAAGACAGAAAAATATGCT 306
    |||||

QY 1706 TGGTTTTCATTTGCTTACCAAAAAACACAAAAAAGTTGCTTTGAGAACTTCA 1765
    |||||
Db 305 TGGTTTTCATTTGCTTACCAAAAAACACAAAAAAGTTGCTTTGAGAACTTCA 246
    |||||

QY 1766 CTGCTCCTATGTGGTACTGCTGAGTCAAAATGTCTATTTTGTCTGTGAAAAATAAATT 1825
    |||||

```

```

Db 245 CCGTCTCCTATGTGGGTACCTGAGTCAAAATGTCTATTTTGTCTGAAAAATAAATT 186
QY 1826 TCCTTCCTTTGTACCAATTTCTGTTTGTATTTTACTAAATCTGTAAATCTGTATTTTCTGT 1885
    |||||
Db 185 TCCTTCCTTTGTACCAATTTCTGTTTGTATTTTACTAAATCTGTAAATCTGTATTTTCTGT 126
    |||||
QY 1886 TTATTCGAAAATTTGATGAACTGCAATCCAAATTTGAAAGTTTGTGTCGACGCTGCTA 1945
    |||||
Db 125 TTATTCGAAAATTTGATGAACTGCAATCCAAATTTGAAAGTTTGTGTCGACGCTGCTA 66
    |||||
QY 1946 GCTTAAATGAATGTGTTCTTATTTGTTTATACATTTATATTAATAAATTTGACATTTTTC 2005
    |||||
Db 65 GCTTAAATGAATGTGTTCTTATTTGTTTATACATTTATATTAATAAATTTGACATTTTTC 6
    |||||
QY 2006 TAAAT 2010
Db 5 TAAAT 1

RESULT 27
CF140548 444 bp mRNA linear EST 06-AUG-2003
LOCUS UI-HF-CB0-apk-b-01-0-UI-r2 NIH MGC_210 Homo sapiens cDNA clone
DEFINITION IMAGE:3098017 5', mRNA sequence.
ACCESSION CF140548
VERSION CF140548.1 GI:33255992
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 444)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
FEATURES
Location/Qualifiers
1..444
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3098017"
/tissue_type="CNCAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC 210"
/notes="Organ: Prostate; Vector: p773 Pac; Site_1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into p773 Pac vector. The library tag
sequence located between the Not I site and the polyA tail
is CCCAC. Tissue was provided by Tim Ratliff."
ORIGIN

```

Query Match 21.0%; Score 421.6; DB 14; Length 444;
 Best Local Similarity 99.1%; Pred. No. 4.2e-74; Indels 0; Gaps 0;
 Matches 424; Conservative 0; Mismatches 4;

QY 1 GGAATAACTGTTCTCTTCTGCGCAGAGAACCTGCTTCAAAGCAGAGTAGCAGTTTC 60
 DB 17 GGAATAACTGTTCTCTTCTGCGCAGAGAACCTGCTTCAAAGCAGAGTAGCAGTTTC 76

QY 61 CGAGTCAGCTGGCTAAATCATCCAGAGAGTAATGCAACCCATGCTTAGAATC 120
 DB 77 CGAGTCAGCTGGCTAAATCATCCAGAGAGTAATGCAACCCATGCTTAGAATC 136

QY 121 GCTGGCTGTTCTTCTGTTGGTGGTGGTAAATGCTGGGACAGTGGCTGTCTAGTGCCT 180
 DB 137 GCTGGCTGTTCTTCTGTTGGTGGTGGTAAATGCTGGGACAGTGGCTGTCTAGTGCCT 196

QY 181 CAGTGGAGAGTGTGGCTCTCAATTGAAAAACAACATCGTGTGTTTGAATACTCTGGGAA 240
 DB 197 CAGTGGAGAGTGTGGCTCTCAATTGAAAAACAACATCGTGTGTTTGAATACTCTGGGAA 256

QY 241 GGACTGTGGATGAATGGCTGAGGAGGCTAAACATCAGGATGAGTGCAGTAAATCTATGAT 300
 DB 257 GGACTGTGGATGAATGGCTGAGGAGGCTAAACATCAGGATGAGTGCAGTAAATCTATGAT 316

QY 301 TCCCTGTGGCTCTTCTCCGGACCTACAGGACCCAGAGAGTGTGTGCTGCTCC 360
 DB 317 TCCCTGTGGCTCTTCTCCGGACCTACAGGACCCAGAGAGTGTGTGCTGCTCC 376

QY 361 GTGATGCTCTTCTGGCTTTCATGATGGCCATCTGGCATGAATGCCAGGTCAGC 420
 DB 377 GTGATGCTCTTCTGGCTTTCATGATGGCCATCTGGCATGAATGCCAGGTCAGC 436

QY 421 GGGGACAA 428
 DB 437 GGGGACAA 444

RESULT 28
 AJ507044
 LOCUS
 DEFINITION Mus musculus BALB/c (Kahlem P) Mus musculus cDNA clone
 127, mRNA sequence.

ACCESSION
 VERSION AJ507044.1 GI:23034129
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 752)

TITLE
 JOURNAL A gene expression map of human chromosome 21 orthologs in the mouse
 COMMENT Nature (2002) In press
 Contact: Kahlem P
 Vertebrate Genomics
 Max-Planck Institute for Molec. Genet.
 Ihnestrass 73, D-14195 Berlin, GERMANY.
 Location/Qualifiers
 1..752
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="127"
 /clone_lib="Mus musculus BALB/c (Kahlem P)"

FEATURES
 source

ORIGIN
 Query Match 20.1%; Score 404.4; DB 9; Length 752;
 Best Local Similarity 82.6%; Pred. No. 1e-70;
 Matches 498; Conservative 0; Mismatches 102; Indels 3; Gaps 3;

QY 106 CATGCTCTAGAAATCGCTGGCTGTTTCTTGGTGGTGGTAAATGCTGGGACAGTGGCT 165
 DB 61 CTGTGCTCTTCAAATGCTGCACTGGTGTCTTGGTGGTGGTAAATGCTGGGACAGTGGCT 120

QY 166 GTCACTGTCTATGCTCAGTGGAGAGTGTGGGCTTTCATTGAAAAACAACATCGTGTGTTT 225
 DB 121 GTCACTATATGCTCAGTGGAGAGTGTCTGCTTTCATGAAAAAGTAACATTTGCTGTGTT 180

QY 226 GAAAACTTCTGGGAAGGACTGTGGATGAATTCGGTGGGACAGCTTAACATCAGGATCGAG 285
 DB 181 GAGAACCGCTGGGAAGGCTTGTGGATGAATTTATGAGGCATGCCAACATCAGAAATCGAG 240

QY 286 TGCAAAATCTATGATTCCTGCTGGCTCTTTCTCCGGACCTACAGGACCCAGAGGACTG 345
 DB 241 TGCAAGGTCTACGACTCCCTGCTGGCTCTTAGTCCAGACCTCCAGGCATCCGAGGACTG 300

QY 346 ATGTGTGCTGCTTCCGCTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAA 405
 DB 301 ATGTGTGCTGCTGCTGCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAA 360

QY 406 TGCAACAGGTGTCACGGGGACAAATGAGAGGTGAAGGCTCACATTTCTGCTACGGCTGGA 465
 DB 361 TGCAACAGGTGTCACGGGGACAAATGAGAGGTGAAGGCTCACATTTCTGCTACGGCTGGA 420

QY 466 ATCATCTTTCATCATCAGGGCATGTGTGGTGTCTATCCCTGTGAGCTGGGTTGCCATGCC 525
 DB 421 ATCATCTTTCATCATCAGGGCATGTGTGGTGTCTATCCCTGTGAGCTGGGTTGCCATGCC 480

QY 526 ATCATCAGAGATTTCTATTAACCTCAATAGTGAATGTG-CCCAAAAACGTGAGCTGGAGA 584
 DB 481 ATCATCAGAGATTTCTATTAACCTCAATAGTGAATGTG-CCCAAAAACGTGAGCTGGAGA 540

QY 585 AGCTCTCTACTTAGATGGACCAACGACACTGTGTGCTGATTTGTGGAGAGCTGTGTTCTG 644
 DB 541 AGCTCTCTACTTAGATGGACCAACGACACTGTGTGCTGATTTGTGGAGAGCTGTGTTCTG 600

QY 645 CTGGTCTTCTGTTGCAACGAAAGAGCAGTAGCTACAGATA-CTCGATACCTTCCCCATC 703
 DB 601 GTGGGTGATTTGTG-CTGAAAGGAGCAACAGTTACAGGTACCTCGGTACCATNCCATC 659

QY 704 GCA 706
 DB 660 GCA 662

RESULT 29
 AW237781/c
 LOCUS
 DEFINITION AW237781 400 bp mRNA linear EST 13-DEC-1999
 xm81e07.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2690628 3',
 mRNA sequence.

ACCESSION
 VERSION AW237781.1 GI:6570170
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 400)

TITLE
 JOURNAL NCI-CGAP
 COMMENT NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/hbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 379.
Location/Qualifiers
1. .400

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2690628"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1509552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 19.9%; Score 399; DB 10; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.4e-69;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1612 AGAAATCATATATGATGATATATTTAATAAGTATTGAGTACAGACTTTGAGGTTTC 1671
Db 400 AGAAATCATATATGATGATATATTTAATAAGTATTGAGTACAGACTTTGAGGTTTC 341
Qy 1672 ATCAATATAATAAAGACAGAGAAATATGCTTGGTTTCATTTGCTTACCAAAAAA 1731
Db 340 ATCAATATAATAAAGACAGAGAAATATGCTTGGTTTCATTTGCTTACCAAAAAA 281
Qy 1732 CAACAACAAAAAAGTTGCTTTGAGAACTTCACCTGCTCCTATGTTGGTACCTGAGTC 1791
Db 280 CAACAACAAAAAAGTTGCTTTGAGAACTTCACCTGCTCCTATGTTGGTACCTGAGTC 221
Qy 1792 AAAATTGTCATTTTGTCTGTAATAATAAATTTCCCTTCTGTAACATTTCTGTTAGT 1851
Db 220 AAAATTGTCATTTTGTCTGTAATAATAAATTTCCCTTCTGTAACATTTCTGTTAGT 161
Qy 1852 TTATCTAAAATCTGTAATACTGTAATTTCTGTTATTTTCAAAATTTGATGAAACTGACA 1911
Db 160 TTATCTAAAATCTGTAATACTGTAATTTCTGTTATTTTCAAAATTTGATGAAACTGACA 101
Qy 1912 ATCCAATTTGAAAGTTTGTGCGACGCTGCTAGCTTAAATGATGTTCTATTGCT 1971
Db 100 ATCCAATTTGAAAGTTTGTGCGACGCTGCTAGCTTAAATGATGTTCTATTGCT 41
Qy 1972 TTATACATTTATTAATAAATGATACATTTTCTTAAT 2010
Db 40 TTATACATTTATTAATAAATGATACATTTTCTTAAT 2

RESULT 30
CK333041
LOCUS CK333041 642 bp mRNA linear EST 19-DEC-2003
DEFINITION H8229C06-5 NIA Mouse Unique Gene Set Version 2 Mus musculus cDNA
clone H8229C06 5', mRNA sequence.
ACCESSION CK333041
VERSION CK333041.1 GI:40232656
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 642)
VanBuren,V., Pao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
Martin,P.R., Staggs,C.A., Bassey,U., Aiba,K., Hamatani,T.,
Kargul,G.J., Luo,A.G., Kelsa,J., Hide,W. and Ko,M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set

Genome Res. 12 (12), 1999-2003 (2002)
22354164
12466305
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: H8229 row: C column: 06
Seq primer: M13 Reverse
High quality sequence stop: 642
POLYA=No.

FEATURES
source

1. .642
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="niaEST:H8229C06-5"
/db_xref="taxon:10090"
/clone="H8229C06"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/clone_lib="NIA Mouse Unique Gene Set Version 2"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This clone is among a rearranged set of 11,424 clones from more than 20 cDNA libraries."

ORIGIN

Query Match 19.8%; Score 398.6; DB 14; Length 642;
Best Local Similarity 83.5%; Pred. No. 1.6e-69;
Matches 452; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 90 GAGGATAATGGCAACCACCTAGAAATCGCTGGGCTGTTCTTGTGGTGGTGAAT 149
Db 102 GAGGACGATGGCAACCACCTAGCTTCAATGGCTGCTGGTGGTGGTGGCAT 161
Qy 150 GGTGGGCAAGTGGCTGTCATGTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 209
Db 162 GGTGGGCAAGTGGCTGTCATGTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 221
Qy 210 CAACATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 269
Db 222 TAACATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 281
Qy 270 TAACATCAGGATCGAGTGCAGAAATCTATGATCCCTGCTGCTCTTTCTCCGGAC 329
Db 282 CAACATCAGAAATGCAAGTCTAGACCTCCCTGCTGCTCTTAGTCCAGACCTCCA 341
Qy 330 GGCAGCCAGGAGTGTGTCGCTTCCGTGATGTCCTTCTTGGCTTTTCATGATGGC 389
Db 342 GGCATCCGAGGACTGATGTGTCGCTGCTGCTTCTTGGCTTTTCATGATGAGC 401
Qy 390 CATCTCTGGCATGAAATGACCCAGGTGCACGGGGGACAAATGAGAAGTGAAGGCTC 449
Db 402 CATCTCTGGAATGAGTGCACCAAGTGCACGGGGGACCAATGAGAAGTGAAGGCTC 461
Qy 450 TTGTGTGACGGTGGAAATCATCTTTCATCATCAGGGGATGTGGTGTCTATCCCTGT 509
Db 462 CTTGTGTGACAGCGGAAATCATCTTTCATCATCAGGGGCTTGGTGTGTCTATCCCT 521
Qy 510 CTGGGTTGCCAATGCAATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCAAA 569
Db 522 CTGGGTTGCCAATGCAATCATCAGAGATTTCTATCAACCACTGTGGATGTGGCCCT 581
Qy 570 ACCTGAGCTTGGAGAGCTCTCTATCTAGGATGGACACCGGACCTGGTGTGATGTTGG 629
Db 582 GCSCGAGCTGGGAGAGCCCTCTACATAGGCTGGACCAACAGCGCTGGTGTGCTGCT 641
Qy 630 A 630

Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:977992
 Seq primer: custom primer used
 High quality sequence stop: 518.
 Location/Qualifiers
 1..551
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1921700"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse kidney mkia"
 /note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
 (CACTGTG); Site 2: DraIII (CACCATG); 1st strand cDNA
 was primed with an oligo(dT) primer
 [ATGTCCTTTTCTTTTCTTTT]; double-stranded cDNA was
 ligated to a DraIII adaptor [TGTGGCTACTGG], digested
 and cloned into distinct DraIII sites of the pME18S-FL3
 vector (5' site CACTGTG, 3' site CACCATG). XhoI should
 be used to isolate the cDNA insert. Size selection was
 performed to exclude fragments <1.5kb. Library
 constructed by Dr. Sumio Sugano (University of Tokyo
 Institute of Medical Science). Custom primers for
 sequencing: 5' end primer CTTCTGCTCTAAAGCTCG and 3' end
 primer CGACCTGCAGCTCGACACA."

ORIGIN

Query Match 19.7%; Score 396.6; DB 9; Length 551;
 Best Local Similarity 83.5%; Pred. No. 4.1e-69; Indels 0; Gaps 0;
 Matches 450; Conservative 0; Mismatches 89;

QY 90 GAGGATAATGGCAACCCATGCTTGAATACTCGTGGCTGTTCTTGTGTGTTGGAAT 149
 DB 13 GAGGACGATGGCAACTACGCTCTTCAATGCTGCACCTGGTGTGTGTTGCAAT 72
 QY 150 GGTGGGCACAGTGGCTGCTCAGTGTATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 209
 DB 73 GGTGGGCACAGTGGCTGCTCAGTGTATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 132
 QY 210 CAACATCGTGGTGTGTTGAAACCTCTGGAGAGCTGTGGATGATTCGGTGGAGCGGC 269
 DB 133 TACATTTGTTGTTTGAGAACCGCTGGAGAGCTGTGGATGATTCGGTGGAGCGGC 192
 QY 270 TAAATCAGGATGACGTGCAAAATCTATGATTCCTGCTGGCTCTTTCTCCGACCTACA 329
 DB 193 CAACATCAGATGCAATGCAAGTCTACGACTCCCTGCTGGCTCTTAGTCCAGACCTCCA 252
 QY 330 GGCAGCCAGAGGACTGATGTGCTGCTTCCGCTGATGCTCTTGGCTTTTCATGATGC 389
 DB 253 GGCATCCCGAGGACTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
 QY 390 CATCTCTGGCATGAAATGCACAGGTGACCGGGGACAACTGAGAGGTGAAGGCTTCACAT 449
 DB 313 CATCTCTGGCATGAAATGCACAGGTGACCGGGGACAACTGAGAGGTGAAGGCTTCACAT 372
 QY 450 TCTGCTGACGGCTGGAATCATCTTCATCATACCGGCTATGGTGTGCTCATCCCTGTGAG 509
 DB 373 CTGCTGACAGCCGGAATCATCTTCTCATCACCGGCTGCTGCTGCTGCTGCTGCTGCTG 432
 QY 510 CTGGTGTGCAATGCCATCATCAGAGTTCTTATTAATCAATAGTGAATGTTGCCAAAA 569
 DB 433 CTGGTGTGCAATTCATCATCAGAGCTTCTAACCACCACTGTTGGATGTGGCCCTAAA 492
 QY 570 ACTGAGCTTGGAGAGCTCTTCTACTTAGATGACCAACCGGCTGCTGCTGCTGCTGCTG 628
 DB 493 GCGGAGCTGGGAGAGCCCTCTACATAGGCTGGACCAACCGGCTGCTGCTGCTGCTGCTG 551

RESULT 33

BI104115
 LOCUS 602890432F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5035638 5',
 DEFINITION mRNA sequence.
 ACCESSION BI104115
 VERSION BI104115.1 GI:14555008
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 754)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1098 Row: p column: 07
 High quality sequence stop: 2
 High quality sequence stop: 653.
 Location/Qualifiers
 1..754
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5035638"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 Stem cell origin.
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 19.4%; Score 390.8; DB 12; Length 754;
 Best Local Similarity 78.3%; Pred. No. 5.5e-68; Indels 3; Gaps 3;
 Matches 506; Conservative 0; Mismatches 137;

QY 90 GAGGATAATGGCAACCCATGCTTGAATACTCGTGGCTGTTCTTGTGTGTTGGAAT 149
 DB 106 GAGGACATGGCAACCTAGCTCTTCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCT 165
 QY 150 GGTGGGCACAGTGGC-TGTCACTGTCATGCTCAG-TGGAGAGTGTGGCTTCATTGAA 207
 DB 166 GGTGGGCACAGTGGCTGTTGACTATTATGCTCAGTGGAGAGTGTGCTTCATCGAA 225
 QY 208 AACCAATCGTCTGTTTGAACCTCTCGGAAGGACTGTGGATGAATTCGTCAGGCAG 267
 DB 226 AGTAACATTTGCTGTTTGAGAACCGCTGGGAGGCTTGTGGATGAATGAGGAT 285
 QY 268 GCTAACATCAGGATCAGTGGCAAAATCTATGATTCCTGCTGCTGCTTTCCTCCGACCTA 327
 DB 286 GCCAATCAGAAATCAGTGGCAAGGCTTACGACTCCCTGCTGCTGCTTCTAGTCCAGACCTC 345
 QY 328 CAGGACGACAGAGGACTGATGTGCTGCTTCCGATGCTCTTCTTGGCTTTCATGATG 387
 DB 346 CAGGATCCAG-GGACTGATGTGCTGCTGCTGCTTCTTGGCTTTCATGACA 404
 QY 388 GCCATCTTTGGCATGAAATGCAACAGGTGCAACGGGAGCAATGAGAGGTGAAGGCTCAC 447
 DB 405 GCCATCTCGGAATGAAGTGCACAGATGCAACGGGAGCAATGAGAGGTGAAGGCTCAC 464

QY 448 ATTCTGCTGAGCGCTGGATCATCTTCATCATCAGCGGCATGGTGGCTCATCCCTGTG 507
 DB 465 ATCTTGCTGAGACGGCGGAATCATCTTCTTCATCAGCGGCTTGGTGGCTCATCCCTGTG 524
 QY 508 AGCTGGGTGCGCAATGCCATCATCAGAGATTCTATAACTCAATAGTGAATGTGCCCAA 567
 DB 525 AGCTGGGTGCGCAATGCCATCATCAGAGATTCTTACAACCCATCGTGGATGGCCCTA 584
 QY 568 AAACTGAGCTTGAGAGAGCTCTCTACTTAGGATGGACACCGGACCTGGTGGCTGATTGT 627
 DB 585 AAGCGGAGCTGGGAGAGCGCTCTACATAGGCTGGACACCGGCTGGTGGCTGATCGCT 644
 QY 628 GGAGGAGCTGTGTTCTGCTGGTGTGTTGTTGTCAGAACGAAAGAGCAGTACGATAC 687
 DB 645 GGAGGAGCAGTGTACAGTAGTGTGTCAGGTGTCTGAAAGGACCAACAGTTACAGGTAC 704
 QY 688 TCGATACCTTCCCATCGCACAAACCCAAAAAGTTATCACACCGGAA 733
 DB 705 TCGGTACCAATCCCATCGCACCAATCAACGGGAGTTACCAACCGCGGAA 750

RESULT 34
 CB455630 580 bp mRNA linear EST 26-MAR-2003
 LOCUS 712787 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION CB455630
 ACCESSION CB455630
 VERSION CB455630.1 GI:29262012
 KEYWORDS EST.
 SOURCE Bos taurus (cow)

ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS Smith,T.P.L., Roberts,A.J., Ehternkamp,S.E., Chitko-McKown,C.G.,
 Wray,J.E. and Keele,J.W.
 TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
 JOURNAL Unpublished (2003)
 COMMENT Contact: Smith TPL
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

EMAIL: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross match v0.990329.
 Plate: FOY8070 row: H column: 23
 Seq primer: GTAATACGACTCATATAGG.
 Location/Qualifiers

FEATURES

source
 1..580
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site_1: EcoRI; Site_2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."

ORIGIN

Query Match 18.2%; Score 365.8; DB 14; Length 580;
 Best Local Similarity 80.4%; Pred. No. 5.9e-63;
 Matches 467; Conservative 0; Mismatches 107; Indels 7; Gaps 3;
 QY 936 TTCACGAGAATGAGATATTAACCAATGCTTTCATGTTCTAGAAAGTATAGTAATTG 995
 DB 1 TTCATCTGCATGACAAATAGTACACATGCTTTCATTTCTATATAAGAACATATTG 60
 QY 936 TTTTCTAAGTGGTTCAGAGCACTACTCTTTTATCATTTTACTTCAAAATGACATGCTA 1055

DB 61 TTTTCAAAATGGTTTCATGTTTATCTCTCTTTTATTAGTTACTTCAAAATGACATGTTG 120
 QY 1056 AAGACTGCATTATTTTACTACTGTAAATTTCTCCACGACATAGCATATGTACATAGATGA 1115
 DB 121 GAGAC--ATTATATTTTACAACCTGTACTTCTTATGACAGCGGTTATGTATATAATGA 178
 QY 1116 GTGTAAACATTATATCTCACATA---GAGACATGCTTATATGGTTTTTATTTAAATGAA 1171
 DB 179 GTGTAACTTCTGTGTACATACGTTGATACAGGCTTATATAGTCCCTATTTTAAATGA 238
 QY 1172 ATGCCAGTCCATTACACTGAATAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGG 1231
 DB 239 ACACCTGATTTCATACACTGAATAATAGAAATTCACCTATTGCTTTTCATGGGAACCAAG 298
 QY 1232 ATAGGGTTGAAGAGGTGTTACTTATTAATGTTTAAAAACAGCTTAGGGATTAAATGTCCTCC 1291
 DB 299 GTAAAGTTGAAGAAAGTTAATTAATTTCTTAAAAACAGCTTAGTGAATGCACCTTA 358
 QY 1292 ATTTATAATGAGATTAAAAATGAAGCTTTTAATCAGCAATGTGAAAGAAATGAATGCT 1351
 DB 359 ATTTATAATGAAAGTTAAAAATGCAAGCTCTAATCAGACGCGTAAATGGAACCTAAATGCT 418
 QY 1352 TTCTGATATGCTGTTTTTTAGCCTAGGAGTTAGAAATCCTACTTCTTTATCCTCTCTC 1411
 DB 419 TTTTGATATCTGTTTTTTCAGCCTATGAGTTAGAAATCCAAATTCCTTTTCTCTCATTC 478
 QY 1412 CCAGAGGCTTTTTTTTCTTGTGTATTAATAAATCAAT-TTTTAAACGACAGATATTTGT 1470
 DB 479 CCAGAGGCTTCTCTTTCTTGTATTAATGGTCATCTTTTAAAGGACAGATATTTGT 538
 QY 1471 CAAGGGCTTTTGCAATCAAACTGCTTTTCCAGGGCTTACT 1511
 DB 539 CAAGGGCTATGCAATCAAACTGCTTTTTCAGAGCTATACT 579

RESULT 35

CB425534 584 bp mRNA linear EST 25-MAR-2003
 LOCUS 600565 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION CB425534
 ACCESSION CB425534
 VERSION CB425534.1 GI:29196672
 KEYWORDS EST.
 SOURCE Bos taurus (cow)

ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS Smith,T.P.L., Roberts,A.J., Ehternkamp,S.E., Chitko-McKown,C.G.,
 Wray,J.E. and Keele,J.W.
 TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
 JOURNAL Unpublished (2003)
 COMMENT Contact: Smith TPL
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

EMAIL: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross match v0.990329.
 Plate: FOY8028 row: H column: 16
 Seq primer: GTAATACGACTCATATAGG.

FEATURES

source
 1..584
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 6BOV"
 /note="Vector: pCDNA3.1; Site_1: EcoRI; Site_2: NotI;

Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."

ORIGIN

Query Match 18.0%; Score 360.8; DB 14; Length 584;
 Best Local Similarity 81.7%; Pred. No. 5.9e-62;
 Matches 478; Conservative 0; Mismatches 99; Indels 9; Gaps 5;
 QY 934 ATTTCAGCAGATGAGATATT-ARACCCAAAGCTTTGATTTCTTAGAAGTATAGTAAT 992
 Db 1 ATTTCAGCTGCATGAGAAATTTAGACATTCGCTTTGATTTATTTCTTAAAGACAGTAAT 60
 QY 993 TTGTTTCTTAAGTGGTTCACAGCATCTACTCTTTTATCAITTAATCTACTTCAAAATGACATTG 1052
 Db 61 TTGTTTCTTAAATGGGTCATGTTTATCTCTCTTTTATTAGTTACTTCAAAATGACATTG 120
 QY 1053 CTAAGACATGCAATTTTACTACTGTATTTCTCCACGATAGCATTATGACATAGA 1112
 Db 121 TGGGAGAC--AATTATTTACAACTGTGACTTCTATGACAGAGCGTATGTATATAGA 178
 QY 1113 TGAGTGTAACTTTATATCTACA----TAGAGACATGCTTATATGTTTATTTTAAAT 1168
 Db 179 TGAGTGTGACATTTCTGTCTACATAGGTAGAGACAGCTTATATAGTCTTATTTTAAAT 238
 QY 1169 GAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAATCA 1228
 Db 239 GAAACACTGATTCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAATCA 298
 QY 1229 TGATAGGTTTCAAGAGGTTACTATTAAATGTTTAAACAGCTTAGGGATTAAATGTC 1288
 Db 299 GGGTTAAGATTGAAGAGGTTAATTAATTTGTTTAAACAGCTTAGGTATGATGATGCAC 358
 QY 1289 TCCATTTAATGAAGAT--TAAATGAAGGCTTTAATCAGCAATTTAAAGGAATTCGAAT 1347
 Db 359 TTAATTTAATGAAGGTTGTAATAAGGAAGCTTTAATCAGCAGGTTAAAGGAATTCGAAT 418
 QY 1348 GCGTTTCTGATATGCTGTTTTTACCTAGGAGTTAGAAATCCCTAACTCTTTATCTCT 1407
 Db 419 GCGTTTCTGATATGCTGTTTTTACGCTTANGAGTTAGAAATCCAAATTCCTTTTCTCA 478
 QY 1408 TCTCCAGAGGCTTTTTTTCTGTTGTTAAATTAACAT--TTTTAAACGAGATATT 1466
 Db 479 TTCCCAAGAGGCTTCCCTTTCTGTTATTAATTAATGAGGCTCTTTTAAAGAGATATT 538
 QY 1467 TTGTCAGGGGCTTTGCAATCAAACTGCTTTTTCAGGCTTACT 1511
 Db 539 TTGTCAGGGGCTTTGCAATTCAACTGCTTTTTCAGAGCTTACT 583

RESULT 36
 BE285566
 LOCUS
 DEFINITION 601095576F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3490218 5', mRNA sequence.
 BE285566
 BE285566.1 GI:9163274
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 918)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM8532 row: g column: 19
 High quality sequence, stop: 629.

FEATURES

Location/Qualifiers
 1..918
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:3490218"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam5"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"

ORIGIN

Query Match 17.9%; Score 360.4; DB 10; Length 918;
 Best Local Similarity 83.0%; Pred. No. 6.5e-62;
 Matches 470; Conservative 0; Mismatches 91; Indels 5; Gaps 5;
 QY 90 GAGGATAATGCAACCCATGCTTAGAAATCGCTGGCTGTTCTTGGTGGTGAAT 149
 Db 74 GAGGACGATGGCAACCTACGCTCTCAATGGCTGCATGGTGGTGGTGGCAT 133
 QY 150 GGTGGGCACTGGCTGTCACCTGTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 209
 Db 134 GGTGGGCACTGGCTGTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAA 193
 QY 210 CAACATCTGTTGTTTTGAAAACTTCTGGGAAGGACTGTGGATGAATTCGCTGAGGAGGC 269
 Db 194 TAACTTGTGTTTGAAGACCGCTGGAAGGCTTGTGGATGAATTTGTATGAGGAGTC 253
 QY 270 TAACT-ATCAGATGAGTG-CAAAATCTATGATTCCTGCTGGCTCTTCTCCGAGCTA 327
 Db 254 CAACATCAGATGAGTGCCCAAGTCTACAGCTCCCTGCTGGCTCTTAGTCCAGACCTC 313
 QY 328 CAGGAGCCAGAGGACTGATGTGTGCTCTTCCGTGATGCTCTTGGTTTCATGATG 387
 Db 314 CAGGATCCCGAGACTGATGTGTGCTGCTGCTCTTGGCTTCTTGGCTTTCATGAG 373
 QY 388 -GCCATCTTGGCATGAAATGCAAGGTGCAAGGAGCAATGAGAGGTGAAGGCTCA 446
 Db 374 AGCCATCTCTCGAATGAAGTGCACAGATGCACGGGGACGATGAGAACGTTGAAGAGCG 433
 QY 447 CATTTCTGCTGAGGCTGGAATCATCTTATCATCACGGGATGTTGTTGCTCATCCCTGT 506
 Db 434 CATCTGCTGACAGCGGAATCATCTTCTTATCACCAGCTTGGTTGTTGCTCATCCCTGT 493
 QY 507 GAGCTGGGTGTCGAATGCTCATCAGAGATTTCTATAACTCAATAGTGAATTTGCCCA 566
 Db 494 CAGCTGGGTGTCGAATTCATCATCAGAGACTTCTACACCCACTGTTGATGTTGGCCT 553
 QY 567 AAAAGCTGAGCTGGAGAGCTCTTACTTAGGATGACACCGGCACTGGTGTGATGT 626
 Db 554 AAAAGCTGAGCTGGAGAGG-CTTACATAGGTGACACACAGCGCT-GTGCTGATGCG 611
 QY 627 TCGAGGAGCTGTTCTGCTGCTGTT 652
 Db 612 TCGATGAGCCCTGTTCTGTTGTTGTT 637

RESULT 37
 CB235084
 LOCUS
 DEFINITION AGENCOURT_11444483 NIH_MGC_166 Mus musculus cDNA clone
 IMAGE:30246099 5', mRNA sequence.
 CB235084
 ACCESSION

```

VERSION CB235084.1 GI:28286662
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: T. Rajendra Kumar and Martin M. Matzuk
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM02 row: k column: 04
High quality sequence stop: 524.
FEATURES
Location/Qualifiers
source
1..767
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30246099"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 166"
/note="Organ: brain; Vector: pDNR-LIB; Site 1: SfiI
(ggcatatggcc); Site 2: SfiI (ggcgctcgcc);
Non-normalized full-length enriched library; 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTAGAGCCGAGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 2.05
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Corp. ."
ORIGIN
Query Match 17.7%; Score 355.6; DB 14; Length 767;
Best Local Similarity 82.0%; Pred. No. 6.2e-61;
Matches 446; Conservative 0; Mismatches 94; Indels 4; Gaps 3;
QY 90 GAGGATATGGCAACCCATGCTTAAATCGCTGGGCTTTCTTGGTGGTGGAT 149
DB 187 GAGGACGATGGCAACCTACGCTTCAATGGCTGCTGGTGGTGGTGGCAT 246
QY 150 GGTGGGACAGTGGCTGCTAGTCTAGTGGAGAGTGTGGCTTCAATGAAA 209
DB 247 GGTGGGACAGTGGCTGCTAGTCTAGTGGAGAGTGTGGCTTCAATGAAA 306
QY 210 CAACATCGTGTGTTTGAACACTTCTGGGAAGAGTGTGGATGAATGGTGGG 269
DB 307 TAAATTTGGTGTGTTGAAGACCGCTGGGAAGGCTTGTGGATGAATGGTGG 366
QY 270 TAAATCGAGTGGTGGCAAACTATGATCCCTGGCTGCTTCTCCGACCTACA 329
DB 367 CAACATCAAGATGAGTGGCAAGGCTTACGACTCCCTGGCTGCTTGTAGTCC 426
QY 330 GCGAGCCAGAGGACTGATGTGCTGCTCCGATGCTCTTCTGGCTTTTATGATGC 389
DB 427 GGCATCCGAGGACTGATGTGCTGCTCCGATGCTCTTCTGGCTTTTATGATGC 486
QY 390 CATCTTGGCATGAATGACCCAGTGCAGCGGGGACATGAGAGGTGAGGCTCACAT 449
DB 487 CATCTTGGCATGAATGACCCAGTGCAGCGGGGACATGAGAGGTGAGGCTCACAT 546
QY 450 TCTGCTGACGGCTGGATCATCTTTCATCATCA-CGGGCATGGTGGTGGCTGTCATCCCTGTGA 508
DB 547 CTGCTGACACCGGATCATCTTTCATCATCA-CGGGCATGGTGGTGGCTGTCATCCCTGTGA 606

```

```

QY 509 GCTGGGTGCCAATGCCATCATCAGAGATTTCTATACTCAATAGTGAATTTGCCAAA 568
DB 607 GCTGGGTGCCAATTTCCATCATCAGAGACTTCTACACCCCACTGGTGGATGGCCCCCTA 666
QY 569 AAGCTGAGCTT--GGAGAGCTTCTTACTAGGATGGACAC--GGCACTGGTGGTGGTGG 625
DB 667 AAGCGGAGCTGGGGAGAGCCCTTACATAGCTGGGACACAGCCGCTGGTGGTGGTGG 726
QY 626 TTGG 629
DB 727 CTGG 730
RESULT 38
BX283446
LOCUS BX283446
DEFINITION BX283446 NIH MGC 75 Homo sapiens cDNA clone IMAGE958091346 ;
ACCSSION BX283446
VERSION BX283446.1 GI:28847900
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
JOURNAL Radehof, U., Schneider, D. and Korn, B.
COMMENT Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rofls
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE958091346.
RZPD; I.M.A.G.E. CDNA Clone Collection;
RZPD; I.M.A.G.E. CDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPD LIB No. 972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rofls
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
Contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMV-M13u, Primer sequence: CGTTGTAACACGACGCGCCAGT.
FEATURES
Location/Qualifiers
source
1..355
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE958091346 ; IMAGE:4606880"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_75"
/note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 17.7%; Score 355; DB 13; Length 355;
Best Local Similarity 100.0%; Pred. No. 9.6e-61;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 ACAGAAACCCCTCTTCAAGCAGAGTAGTAGTCCGAGTCCAGTGGTGGCTTAAACTCA 84

```

Db 1 ACAGAGAACCCCTGCTTCAAAGCAGAGTAGCAGTTCCGGAGTCCAGCTGGCTAAACCTCA 60
 QY 85 TCCACAGAGATATGCAACCCATCGCTAGAAATCGCTGGCTGTTCTTGTGTGGTCTT 144
 Db 61 TCCACAGAGATATGCAACCCATCGCTAGAAATCGCTGGCTGTTCTTGTGTGGTCTT 120
 QY 145 GGAATGGTGGGACAGTGGCTGCTCACTGTCATGCTCAGTGGAGAGTGTGGGCTTCATT 204
 Db 121 GGAATGGTGGGACAGTGGCTGCTCACTGTCATGCTCAGTGGAGAGTGTGGGCTTCATT 180
 QY 205 GAAACCAACATCGTGGTGTGTTTGAACACTTCTGGGAAGACTGTGAGTAATGGTGAGG 264
 Db 181 GAAACCAACATCGTGGTGTGTTTGAACACTTCTGGGAAGACTGTGAGTAATGGTGAGG 240
 QY 265 CAGGCTAACATCAGGATGAGTGCAGAAATCTATGATTCCTGCTGGCTTCTTCCGGAC 324
 Db 241 CAGGCTAACATCAGGATGAGTGCAGAAATCTATGATTCCTGCTGGCTTCTTCCGGAC 300
 QY 325 CTACAGGACCCAGAGGACTGATGTGCTGCTCCGTCATGTCCTTCTTGGCTT 379
 Db 301 CTACAGGACCCAGAGGACTGATGTGCTGCTCCGTCATGTCCTTCTTGGCTT 355

RESULT 39

BE172635
 LOCUS 360 bp mRNA linear EST 21-JUN-2000
 DEFINITION MR0-HT0559-050400-012-e04 HT0559 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE172635
 VERSION BE172635.1 GI:8635361
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 360)
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-MR0-HT0559-050400-012-e04&t3=2000-04-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 15
 High quality sequence stop: 360.

FEATURES

source

1..360
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0559"
 /note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of

ORIGIN

Query Match 17.5%; Score 351; DB 10; Length 360;
 Best Local Similarity 98.6%; Pred. No. 6.1e-60;
 Matches 354; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1082 TTCTCCACGACATAGCATTATGTACATAGATGAGTGTAAACATTATATCTCACATAGAG 1141
 Db 1 TTCTACACGGCATAGCAATTATGTACATAGATGAGTGTAAACATTATATCTCACATAGAG 60
 QY 1142 ACATGCTTATATGTTTTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGAA 1201
 Db 61 ACATGCTTATATGTTTTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGAA 120
 QY 1202 CTCACACTATTCCTTTTCAGGGAATCATGATAGGTTGAGAGGTTACTATTAAATGT 1261
 Db 121 CTCACACTATTCCTTTTCAGGGAATCATGATAGGTTGAGAGGTTACTATTAAATGT 180
 QY 1262 TTTAAAAACAGCTTAGGGATTAAATGTCCTCCATTATATGAAGATTAAAAATGAAGGCTTT 1321
 Db 181 TTTAAAAACAGCTTAGGGATTAAATGTCCTCCATTATATGAAGATTAAAAATGAAGGCTTT 240
 QY 1322 ATACGCAATTGTAAAGGAATTAATGCTTCTGATATGCTGTTTTTAGCCTAGGAGT 1381
 Db 241 ATACGCAATTGTAAAGGAATTAATGCTTCTGATATGCTGTTTTTAGCCTAGGAGT 300
 QY 1382 TAGAAATCCTAACTTCTTTATCTCTCTCCAGAGGCTTTTTTTTCTGTGTATTAA 1440
 Db 301 TAGAAATCCTAACTTCTTTATCTCTCTCCAGAGGCTTTTTTTTCTGTGTAGAA 359

RESULT 40

BE000282
 LOCUS 352 bp mRNA linear EST 05-JUN-2000
 DEFINITION MR0-BN0070-260400-017-e10 BN0070 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE000282
 VERSION BE000282.1 GI:8260515
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 352)
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.G.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-MR0-BN0070-260400-017-e10&t3=2000-04-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 7
 High quality sequence stop: 352.
 Location/Qualifiers
 1..352

TITLE

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

FEATURES

source

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 465)

REFERENCE
AUTHORS
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE
JOURNAL
COMMENT
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouse@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:916450
 Seq primer: custom primer used
 High quality sequence stop: 457.

FEATURES
source
 1..465
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1432382"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse kidney mkia"
 /note="Organ: Kidney; Vector: pME18S-FL3; Site 1: DraIII (CACATGTG); Site 2: DraIII (CACATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGGCGCTTTTATTTTATTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTGTCCTACTGG], digested and cloned into distinct draIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGACCA."

ORIGIN
 Query Match 16.4%; Score 330.6; DB 9; Length 465;
 Best Local Similarity 81.9%; Pred. No. 78-56; Indels 0; Gaps 0;
 Matches 381; Conservative 0; Mismatches 84;

103 ACCCATGCTTTAGAAATCGCTGGCTGTTTCTTGTGGTGTGGTAAATGGTGGCAGATG 162
 1 ACCTACGCTCTTCAATGCTGCTCGGTTGTTGGCGTTGGCATGTTGGCAGCTTG 60

163 GCTGTCATGTCATGCTCAGTGGAGATGTGCGGCTTCATTTGAAACCAATGCTGTT 222
 61 TCTGTTTCTATCATGCTCAGTGCAGCGCTGCTGCTTTCATCGAAAGTAAATGTTGGT 120

223 TTTGAAACCTTTCGGAAGACTGTGGATGTGGTGGAGCGGCTTAAACATCAGCATG 282
 121 TTTGAGAACCGCTGGAGGCTTTGGATGATTTGATGAGCGATGCCAATCAGATG 180

283 CAGTGCAAAATCTATGATTCCTGCTGGCTCTTTCTCGGACCTACAGGCGCCAGAGGA 342
 181 CAGTGCAAGGCTTACGACTCCCTGCTGGCTCTTAGTCCAGACCTCCAGGCATCCCGAGGA 240

343 CTGATGTGCTCTTCCGCTGATGCTCTTCTTGGCTTTTCATGATGCGCATCTTGGCATG 402
 241 CTGATGTGCTCGCTGCTGCTGCTTCTTCTTGGCTTTTCATGACAGCCATCTCGGAATG 300

403 AAATGCACCAGGTGCACGGGGGCAATGAGAGGTGAAGGCTTCAATTTCTGCTGACGGCT 462

301 AAGTGACCAGATGACGGGGGACGATGAGAACTGAAGAGCGCATCTTGTGACAGCC 360
 463 GGAATCATCTTTCATCATCACGGGATGGTGTGCTCATCCCTGTGAGCTGGTTGCCAAT 522
 361 GGAATCATCTTCTTTCATCACGGGTTGGTGTGCTCATCCCTGTGAGCTGGTTGCCAAT 420

523 GCATCATCAGAGATTCTATAACTCAATAGTAGTAATGTTCGCCAA 567
 421 TCCATCATCAGAGACTTCTACACCCACTGGTGAATGTGGCCCTA 465

RESULT 46
LOCUS
 AI547146 771 bp mRNA linear EST 09-AUG-1999
 PN3.1.01.G02.r mynorm Homo sapiens cDNA 5', mRNA sequence.
ACCESSION
 AI547146 GI:4464634
VERSION
 AI547146.1
KEYWORDS
 EST.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
REFERENCE
AUTHORS
 Huang, G.M., Ng, W.L., Farkas, J., He, L., Liang, H.A., Gordon, D., Yu, J. and Hood, L.
TITLE
 Prostate cancer expression profiling by cDNA sequencing analysis
JOURNAL
 Genomics 59 (2), 178-186 (1999)
MEDLINE
 99339982
PUBMED
 10409429
COMMENT
 Contact: Guyang Matthew Huang
 Leroy Hood
 University of Washington
 Department of Molecular Biotechnology, Box 357730, University of Washington, Seattle, WA 98195
 Tel: 5106280100
 Fax: 5106280108
 Email: huanggm@yahoo.com.

FEATURES
source
 1..771
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="mynorm"
 /note="Organ: Prostate; Vector: pBluescript; Directional cDNA library was constructed using Lambda ZP II kit (Stratagene). mRNA was extracted from a frozen normal prostate tissue (Mayo Clinic)."

ORIGIN
 Query Match 16.3%; Score 328.2; DB 9; Length 771;
 Best Local Similarity 70.9%; Pred. No. 1.9e-55; Indels 42; Gaps 5;
 Matches 497; Conservative 0; Mismatches 162;

781 ATGTTTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAATCTATATTACTTTCTCA 840
 13 AGGTTTTTTTAACTTTTACTATAAAGCCATGCAAAATGACAAAATCTATATTACTTTCTCA 72

841 AAATGAGACCCCAAGAAACTTTGATTTACTGTTCTTAAGTCTTAATCTTAATACAGGA 900
 73 AAATGAGACCCCAAGAAACTTTGATTTACTGTTCTTAAGTCTTAATCTTAATACAGGA 132

901 ACTGTGCATCAGCTATTTATGATTTCTATAAGCTATTTTCAGCAGATGAGATTTAAACCC 960
 133 ACTGTGCATCAGCTATTTATGATTTCTTAAGCTATTTTCAGCAGATGAGATTTAAATCC 192

961 AATGCTTTGATTTGTTCTGAAAGTATAGTAATTTGTTTTTCTAAGGTGTTCAAGCACTTA 1020
 193 AATGCTTTGATTTGTTCTGAAAGTATAGTAATTTGTTTTTCTAAGGTGTTCAAGCACTTA 252

1021 CTTCTTTTATCATTTTACTTCAAAATGACATTCGCTAAAGACTGCATTTATTTTACTACTGTA 1080
 253 CTTCTTTTATCATTTTACTTCAAAATGACATTCGCTAAAGACTGCATTTATTTTACTACTGTA 312


```

Db 445 GTGGGTGCTTATCCATCATCAGAAATCTTAGAACCCACTGGCGGATATGGTCCTAAA 504
QY 570 ACCTGAGCTTGAGAGACTCTACTTAGTAGTGACACCGCACTGGTGCCTGATTGTGG 629
Db 505 GCGCCAGCTGGGAGAGCCCTCTACATAGCTGGACCAATCGCTGGGCTGATCGGTTG 564
QY 630 AGAGCTCTGTTCTGCTGCTGTTTTTTTGTTCACGAAAGAGCAGTAGCTACAGATCTC 689
Db 565 AAGAGCACTGTTCTGAAGAGGGGGTGTGCGAATTTGAAGAGGAGCAAAAATTTCTGGTACTT 624
QY 690 GATACCTTCCCATCGCAC 707
Db 625 GGGACCAATCCATGGAC 642

RESULT 48
AY400666
LOCUS 325 bp DNA linear GSS 15-DEC-2003
DEFINITION Pan troglodytes CLDN8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY400666
VERSION AY400666.1 GI:39756655
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 325)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 325)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..325
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>325
/genes="CLDN8"
/locus_tag="HCM0633"
gene
Query Match 16.1%; Score 323.4; DB 29; Length 325;
Best Local Similarity 99.7%; Pred. No. 2.1e-54;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 450 TCTGCTGACGCTGGAATCATCTTCATCATCATACGGGCATGGTGGTCTCATCCCTGTGAG 509
Db 1 TCTGCTGACGCTGGAATCATCTTCATCATCATACGGGCATGGTGGTCTCATCCCTGTGAG 60
QY 510 CTGGGTTGCCAATGCCATCATCAGAGATTCTTAATCAATAGTAGTAATGTGCCAAA 569
Db 61 CTGGGTTGCCAATGCCATCATCAGAGATTCTTAATCAATAGTAGTAATGTGCCAAA 120
QY 570 ACCTGAGCTTGGAGAGCTCTTACTTAGGATGGACCAAGCAGCTGGTGCCTGATTGTGG 629
Db 121 ACCTGAGCTTGGAGAGCTCTTACTTAGGATGGACCAAGCAGCTGGTGCCTGATTGTGG 180

```

```

QY 630 AGGAGCTCTGTTCTGCTCGTTTTTTTGTTCACGAAAGAGCAGTAGCTACAGATCTC 689
Db 181 AGGAGCTCTGTTCTGCTCGTTTTTTTGTTCACGAAAGAGCAGTAGCTACAGATCTC 240
QY 690 GATACCTTCCCATCGCACACCCCAAAAGTTATCACACCGAAAGAGTCAACCGAGCGT 749
Db 241 GATACCTTCCCATCGCACACCCCAAAAGTTATCACACCGAAAGAGTCAACCGAGCGT 300
QY 750 CTACTCCAGAGTCAGTAGTGTAG 774
Db 301 CTACTCCAGAGTCAGTAGTGTAG 325

RESULT 49
BF151294
LOCUS 578 bp mRNA linear EST 29-DEC-2000
DEFINITION uz14e04.y1 NCI CGAP Mams Mus musculus cDNA clone IMAGE:3669054 5',
similar to SW:CLD8_MOUSE Q9Z260 CLAUDIN-8. ;, mRNA sequence.
ACCESSION BF151294
VERSION BF151294.1 GI:11032689
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 578)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
MGI:1429822
Seq primer: -40RP from Gibco
High quality sequence stop: 409.
Location/Qualifiers
1..578
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/strain="C57BL/6J"
/clone="IMAGE:3669054"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

Query Match 15.9%; Score 319.8; DB 10; Length 578;
Best Local Similarity 81.0%; Pred. No. 9.7e-54;
Matches 372; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 90 GAGGATAATGGCAACCCATGCCTTAGAATCGCTGGGCTGTTCTTGGTGGTGTGAAT 149
Db 119 GAGGACGATGGCAACCTACGCTCTCAATGGCTGCATGGTGTGTTGGCAT 178
QY 150 GTGGGCAAGTGGCTGTCTACGTTCATGCCTCAGTGAGAGTGTGGGCTTCATGAAA 209
Db 179 GTGGGCAAGTGGCTGTCTACGTTCATGCCTCAGTGAGAGTGTGCTTCATCGAAG 238

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 3, 2004, 01:55:56; Search time 2675 Seconds
(without alignments)
3645.676 Million cell updates/sec

Title: US-10-063-732-120
Perfect score: 1172
Sequence: 1 MATHALEIAGLFLGGVGMVG.....OKSYHTGKSPSVYRSQYV 225

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10063732/runat_01092004_155516_20440/app_query.fasta_1.391
-DB=GenEmbl -CPWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMF=pcp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063732@cgn_1_1_2527@runat_01092004_155516_20440 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fur.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1172	100.0	1700	9	BC058004 Homo sapi
2	1172	100.0	1835	9	BC020866 Homo sapi
3	1172	100.0	1931	9	HSA250711 Homo sapi
4	1172	100.0	2010	6	AX092388 Sequence
5	1172	100.0	2010	6	AX376290 Sequence
6	1172	100.0	2010	6	AX697259 Sequence
7	1172	100.0	2010	9	AX358707 Homo sapi
8	1172	100.0	81583	2	AC131927 Homo sapi
9	1172	100.0	137955	9	AF000884 Homo sapi
10	1172	100.0	340000	9	AF001707 Homo sapi
11	1167	99.6	176580	9	BS000178 Pan trogl
12	1167	99.6	267172	9	BS000177 Pan trogl
13	1140	97.3	837	6	AX866683 Sequence
14	1140	97.3	837	6	BD146745 Primer fo
15	1134	96.8	191923	2	AP001846 Homo sapi
16	1127	96.2	1890	6	AX879004 Sequence
17	1127	96.2	1890	6	BD157589 Primer fo
18	1127	96.2	1890	9	AK022269 Homo sapi
19	1106	94.4	1002	6	AX497202 Sequence
20	1018	86.9	240479	2	AC096377 Rattus no
21	1012	86.3	678	6	E31609 Tight junct
22	1012	86.3	678	10	AF087826 Mus muscu
23	1012	86.3	2325	10	BC003868 Mus muscu
24	1012	86.3	217278	2	AC113180 Mus muscu
25	1012	86.3	256720	10	AC110241 Mus muscu
26	696	59.4	675	9	HSA250712 Sequence
27	696	59.4	833	6	AX497204 Sequence
28	676	57.7	1160	6	AX697263 Sequence
29	676	57.7	1160	9	AX358094 Homo sapi
30	663	56.6	615	6	AR414198 Sequence
31	663	56.6	615	6	BD109751 EST and e
32	542.5	46.3	902	5	BC059750 Silurana
33	538	45.9	916	5	AF359435 Xenopus l
34	538	45.9	925	5	AB072909 Xenopus l
35	538	45.9	929	5	AB072908 Xenopus l
36	538	45.9	1736	5	AF334677 Gallus ga
37	534	45.6	660	6	BD076388 Human pro
38	534	45.6	1250	6	AX302529 Sequence
39	534	45.6	1250	9	AB000714 Homo sapi
40	534	45.6	1279	6	BD076398 Human pro
41	534	45.6	1294	6	BC016056 Homo sapi
42	534	45.6	1601	9	AF007189 Homo sapi
43	534	45.6	148689	9	AC093168 Homo sapi
44	533	45.5	895	5	AF224712 Xenopus l
45	528	45.1	857	5	XLAJ9313 Xenopus l

ALIGNMENTS

RESULT 1

BC058004	1700 bp	mrna	linear	PRI 07-OCT-2003
LOCUS	Homo sapiens claudin 8, mRNA (cDNA clone MGC:61830 IMAGE:4606880), complete cds.			
DEFINITION	BC058004			
ACCESSION	BC058004.1 GI:34783654			
VERSION	MGC.			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1700)			
AUTHORS	Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Rana,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.G., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
MEDLINE	22388257			
PUBMED	12477932			
REFERENCE	2 (bases 1 to 1700)			
AUTHORS	Strausberg,R.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk Email: cgabs-rcmail.nih.gov Tissue Procurement: CLONTECH cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@hgrl.nih.gov Aktef,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Green,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghghi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lari,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vost,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.			
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IPAL Plate: 48 Row: d Column: 13 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21314655. Location/Qualifiers 1..1700 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:61830 IMAGE:4606880"			
source				

/tissue_type="Kidney"	
/clone_lib="NIH MGC_75"	
/lab_host="DH10B"	
/note="Vector: pDNR-LIB"	
1..1700	
/gene="CLDN8"	
/db_xref="LocusID:9073"	
76..753	
/codon_start=1	
/product="CLDN8 protein"	
/protein_id="AAH58004.1"	
/db_xref="GI:34783655"	
/db_xref="LocusID:9073"	
/translation="MATHALEIAGFLGGVMGVAVTVMPQWRVSFAFIENIVVFE NFEGVMNCVQANIRMOCKIYDSLLALSPDQAARGLCAASVMSFLAMMAILGM KCTRCQDNKVKAHILLPAGIFITGMVVLIPVSWANAIIRDYNSIVYVAQRRE LGKALYLGNTALVLIIVGALFCCVFCNKKSSSYRYSIPSHRTIQKSYHTKKSPSV YSRQYV"	
misc_feature	
160..621	
/note="PMP22_Claudin; Region: PMP-22/EMP/MP20/Claudin family"	
/db_xref="CDD:pfam00822"	
ORIGIN	
Alignment Scores:	
Pred. No.: 5,53e-123 Length: 1700	
Score: 1172.00 Matches: 225	
Percent Similarity: 100.00% Conservative: 0	
Best Local Similarity: 100.00% Mismatches: 0	
Query Match: 100.00% Indels: 0	
DB: 9 Gaps: 0	
US-10-063-732-120 (1-225) x BC058004 (1-1700)	
Qy 1 MetAlaThrHisAlaLeuGluLurleAlaGlyLeuPheLeuGlyGlyValGlyMetValGly 20	
Db 76 ATGGCAACCCATGCCCTAGAAATCGCTGGGCTGTTTCTTGGTGGTGGTGGG 135	
Qy 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheileGluAsnAsnile 40	
Db 136 ACAGTGGCTGCTACTGTCTATGCTCAGTGGAGAGTGTGGCCCTTCATTGAAACACATC 195	
Qy 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnile 60	
Db 196 GTGGTTTTTGAACCTCTGGGAAGACGTGGATGAATGGTGAGGACGCTAACATC 255	
Qy 61 ArgMetClnCysGlySyleTyraSpSerLeuAlaLeuSerProAspLeuGlnAlaAla 80	
Db 256 AGGATGCAAGTGCAAAATCTATGATTCCTCGCTGGCTCTTCTCCGAGCTACAGGACGCC 315	
Qy 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaileLeu 100	
Db 316 AGAGGACTGATGCTGCTCTCCGTGATGTCCTCTTGGCTTTCATGATGCCATCCT 375	
Qy 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisileLeuLeu 120	
Db 376 GGCATGAATGCAACAGGTGCACGGGGGCAATAGAGAGGTGAAGGCTCACATTCGTGTG 435	
Qy 121 ThrAlaGlyIlelePheleleThrGlyMetValValLeuileProValSerTrpVal 140	
Db 436 ACGGCTGGAATCATCTTCATCATCAGGCGCATGGTGGTGTCTCATCCTGTGAGCTGGGT 495	
Qy 141 AlaAsnAlaIleleArgAspPheTyraSerIleValAsnValAlaGlnLysArgGlu 160	
Db 496 GCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAAAACGTGAG 555	
Qy 161 LeuGlyGluAlaLeuTyraLeuGlyTrpThrAlaLeuValLeuileValGlyGlyAla 180	
Db 556 CTTGGAGAGCTCTCTACTAGGATGGACACGGACCTGGTGTCTGATTTGTTGGAGAGCT 615	
Qy 181 LeuPheCysCysValPheCysCysAsnGluLysSerSerSerTyraTyraSerIlePro 200	
Db 616 CTGTTCTGCTCGCTTTTGTTCACACGAAAGAGCAGTAGCTACAGATCTCGATACCT 675	

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:24067 IMAGE:4594155"
/tissue_type="Kidney"
/clone_lib="NIH MGC_75"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
1..1935
/genes="CLDN8"
/db_xref="LocusID:9073"
110..787
/codon_start=1
/product="claudin 8"
/protein_id="AAH20866.1"
/db_xref="GI:18089190"
/db_xref="LocusID:9073"
/translation="NATHALETAGLFLGCGMGTAVTVMPOKRVSAFTENNIVVPE
NFVEGWMNCVQANIRMOCKIYDSELLASPDQAQKGLMCAASVMSFLAFMAILGM
KCTRCGDKKRAKHLITAGIIFITGMVLI PVSWMANAIIRDVFNISVNVQAQKE
LGBALYLGWTTALVITVGALPCFVCCNKKSSSVYSIPSHRTTQKSYHTGKKSPSV
YRSQVY"
misc_feature
194..655
/notes="PMP22_Claudin; Region: PMP-22/EMP/MP20/Claudin
family"
/db_xref="CDD:pfam00822"

ORIGIN
Alignment Scores:
Pred. No.: 6,09e-123 Length: 1835
Score: 1172.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-063-732-120 (1-225) x BC020866 (1-1835)
QY 1 MetalathrHisAlaLeuGluLeuAlaGlyLeuPheLeuGlyGlyValGlyMetValGly 20
DB 110 ATGGCAACCCATCGCTTAGAAATCGCTGGCGCTGTTCTTGTGGTGGTGGTGGTGGC 169
QY 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnAsnIle 40
DB 170 ACAGTGGCTGCTCACTGTCATGCTCAGTCGAGAGTGTGGGCTTTCATTGAAACCAACATC 229
QY 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle 60
DB 230 GTGGCTTTTGGAAACTTCCTGGGAAGACACTGTGGATGAATTCGCTGAGCAGCGCTAACATC 289
QY 61 ArgMetGlnCysIysIleTyrAspSerLeuLeuAlaLeuSerProAspLeuGlnAlaAla 80
DB 290 AGGATGCAGTGCAAAATCATGATTCCTGCTGCTCTTCTCGGACCTACAGCGACGCC 349
QY 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
DB 350 AGAGGACTGATGTGTGCTGCTCCGTGATGTGCTCTTCTGGCTTTTCATGATGCCATCCTT 409
QY 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
DB 410 GGCATGAATTCACCCAGGTGCACGGGGGACATGAGAGGTGAGGCTCACATTCGCTG 469
QY 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
DB 470 ACGGCTGGAATCATCTTCATCATCATCGGCATCGGTGGTGGCTCATCCCTGTGAGCTGGGTT 529
QY 141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
DB 530 GCCAATGCCATCATCAGAGATTCTCTATACTCAATAGTGAATGTGCCCAAAACGTGAG 589
QY 161 LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla 180
DB 590 CTTGGAGAAGCTCTTACTTAGATGGACACCGGACCTGGTGTGCTGATTTGGAGAGACT 649

```

```

FEATURES
  source
    Location/Qualifiers
      1. 1835
        /organism="Homo sapiens"

```

QY 181 LeuPheCysValPheCysAysnGluLysSerSerTyrArgTyrSerIlePro 200
Db 650 CTGTTCTGCTGGTTTTTTGTTGCAACGAAAGAGCAGTAGCTACAGATACCTGATACCT 709
QY 201 SerHisArgThrGlnGlySerTyrHisThrGlyLysSerProSerValTyrSer 220
Db 710 TCCCATCGCAACCAACCAAAAGTTATCACCGGAAAGAGTACCGAGCGTCTACTCC 769
QY 221 ArgSerGlnTyrVal 225
Db 770 AGAAGTCAGTATGTG 784

RESULT 3
HSA250711 1931 bp DNA linear PRI 11-NOV-1999
LOCUS Homo sapiens CLDN8 gene for claudin-8.
DEFINITION AJ250711
ACCESSION AJ250711
VERSION AJ250711.1 GI:6433859
KEYWORDS claudin-8; CLDN8 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Keen, T.J. and Inglehearn, C.F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1931)
AUTHORS Keen, T.J.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1999) Keen T.J., Molecular Medicine Unit,
University of Leeds, Clinical Sciences Building, St James's
Hospital, Leeds, LS9 7TF, UNITED KINGDOM

FEATURES
source
1..1931
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
23..700
/gene="CLDN8"
23..700
/gene="CLDN8"
/codon_start=1
/product="claudin-8"
/protein_id="CA650615.1"
/db_xref="GI:6433860"
/db_xref="GOA:P56748"
/db_xref="SWISS-PROT:P56748"
/translation="MATHALEIAGLFLGGVGMVGTVAVTMPQWRVSFIENNIIVVFE
NFTGCLGNKVCROINRMCKIYDSLIALSPDQAARGLMCAASVMSFLAFNWLIGM
KCTGCTDNKCVKHAHLLTAGIIFITGMVLPVSNVANAIIIRDPSNVNVAOKRE
LGEALYLGTWLTALVLIYGGALFCCVCCNEKSSSYRYSIPSHRTTQKSYHTCKKSPSV
YRSQIV"
polyA_signal
1913..1918
ORIGIN
1913..1918

Alignment Scores:
Pred. No.: 6, 5e-123 Length: 1931
Score: 1172.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-063-732-120 (1-225) x HSA250711 (1-1931)

QY 1 MetAlaThrHisAlaLeuGluIleAlaGlyLeuPheLeuGlyValGlyMetValGly 20
Db 23 ATGGCAACCCATGCTTAGAATCGCTGGGCTGTTCTTGGTGGTGTGGAATGGTGGC 82
QY 21 ThrValAlaValThrValMetProGlnTyrArgValSerAlaPheIleGluLeuAsnIle 40

Db 83 ACAAGTGGCTGTCACTGTCACTGCTCAGTGGAGAGTGTGCGCCTTCATTGAAACCAACATC 142
QY 41 ValValPheGluAsnPheTyrPheGluGlyLeuTyrMetAsnCysValArgGlnAlaAsnIle 60
Db 143 GTGGTTTTTGAAACCTTCTGGGAAGACTGTGGATGATTTGGTGGAGGAGCTAACATC 202
QY 61 ArgMetGlnCysLysIleTyrAspSerLeuLeuAlaLeuSerProAspLeuGlnAla 80
Db 203 AGGATGCAAGTCAAAATCTATGATTCCTCGCTGGCTTTCTCCGGACTACAGCAGCC 262
QY 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
Db 263 AGAGGACTGATGTGTGCTCTCCGTGATGTCTTCTTGGCTTTTCATGATGGCCATCCT 322
QY 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
Db 323 GGCATGAATGCACACAGGTGCACGGGGACAATGAGAAGGTGAGGCTCACATTTCTGCTG 382
QY 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
Db 383 ACAGTGAATCATCTTTCATCATCAGCGCATGGTGGTCTCATCCCTGTGAGCTGGGTT 442
QY 141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
Db 443 GCCAATGCCATCATCAGAGATTTCTATACTCAATAGTAGATGTTGCCAAAAACGTGAG 502
QY 161 LeuGlyGluAlaLeuTyrLeuGlyTyrThrThrAlaLeuValLeuIleValGlyValAla 180
Db 503 CTTGGAGAGACTCTCTACTTAGATGACCGGACCTGGTGTCTGATGATTGTGGAGAGCT 562
QY 181 LeuPheCysValPheCysAysnGluLysSerSerTyrArgTyrSerIlePro 200
Db 563 CTGTTCTGCTGGCTTTTGTGTCACGAAAGAGACGAGTAGCTACAGATCTCGATACCT 622
QY 201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSerProSerValTyrSer 220
Db 623 TCCCATCGCACAAACCAACCAAAAGTTATCACCGGAAAGAGTACCGAGCGTCTACTCC 682
QY 221 ArgSerGlnTyrVal 225
Db 683 AGAAGTCAGTATGTG 697

RESULT 4
AX092388 2010 bp DNA linear PAT 21-MAR-2001
LOCUS Sequence 119 from Patent WO0116318.
DEFINITION AX092388
ACCESSION AX092388
VERSION AX092388.1 GI:13444508
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Eaton, D.L., Filvaroff, E., Gerritsen, M.E., Goddard, A.,
Godowski, P.J., Grimaldi, C.J., Gurney, A.L., Watanabe, C.K. and
Wood, W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 119 08-MAR-2001;
Genentech, Inc. (US)
FEATURES Location/Qualifiers
source
1..2010
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6, 83e-123 Length: 2010
Score: 1172.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0


```

Query Match: 100.00%      Indels: 0
DB:          6            Gaps: 0
US-10-063-732-120 (1-225) x AX092388 (1-2010)

QY 1 MetAlaThrHisAlaLeuGluLeuAlaGlyLeuPheLeuGlyGlyValGlyMetValGly 20
DB 97 ATGGCAACCATGCTTGAATACTGCTGGCTGTTCTTGGTGGTGTGGAAATGGTGGC 156

QY 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnAsnIle 40
DB 157 ACAGTGGCTGTCACCTGTCATGCTCAGTGGAGAGTGTGGCTTCAATGAAACACATC 216

QY 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle 60
DB 217 GTGGTTTTTGAATACTTCTGGAGAGACTGTGGATGAATTCGTGGAGCGCTAACATC 276

QY 61 ArgMetGlnCysLysIleTyrAspSerLeuAlaLeuSerProAspLeuGlnAlaAla 80
DB 277 AGAGTGCAGTGCATAATCTATGATTCCTCGTGTGCTTCTTGGCTTTCATGATGGCCATCCT 336

QY 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
DB 337 AGAGGACTGATGTGCTGCTTCCTCGTGTGCTTCTTGGCTTTCATGATGGCCATCCT 396

QY 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
DB 397 GGCATGAATGCACAGGTGCACGGGGACCAATGAGAGGTGAAGGCTCACATTCCTCTG 456

QY 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
DB 457 ACGGCTGGATCATCTTATCAATCAGGGCATGGTGGTGTCTTCTTGGCTTTCATGATGGCCATCCT 516

QY 141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
DB 517 GCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTGGCAAAACGTGAG 576

QY 161 LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla 180
DB 577 CTGGAGAGCTCTCTACTTAGGATGGACACCGCAGCTGGTGTGCTGATTGTGGAGGAGCT 636

QY 181 LeuPheCysValPheCysAsnGluLysSerSerTyrArgTyrSerIlePro 200
DB 637 CTGTTCTGCTGGTTTTTGTGCAACGAAAGAGCAGTACAGTACAGATACCTGATACCT 696

QY 201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSerProSerValTyrSer 220
DB 697 TCCCATCGCACCAACCAAAAGTTATCACACCGGAAAGAGTACCAGGCGTCTACTCC 756

QY 221 ArgSerGlnTyrVal 225
DB 757 AGAAGTCAGTATGTG 771

RESULT 5
AX376290
LOCUS
DEFINITION
Sequence 357 from Patent WO0168848.
ACCESSION
AX376290.1 GI:19170536
VERSION
AX376290.1
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J.,
Gurney, A.L., Fan, J., Smith, V., Watanabe, C.K., Wood, W.I. and
Zhang, Z.
TITLE
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL
Patent: WO 0168848-A 357 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
Location/Qualifiers

```

```

source
1. .2010
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 6,83e-123      Length: 2010
Score: 1172.00           Matches: 225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 6                    Gaps: 0

US-10-063-732-120 (1-225) x AX376290 (1-2010)

QY 1 MetAlaThrHisAlaLeuGluLeuAlaGlyLeuPheLeuGlyGlyValGlyMetValGly 20
DB 97 ATGGCAACCATGCTTGAATACTGCTGGCTGTTCTTGGTGGTGTGGAAATGGTGGC 156

QY 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnAsnIle 40
DB 157 ACAGTGGCTGTCACCTGTCATGCTCAGTGGAGAGTGTGGCTTCAATGAAACACATC 216

QY 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle 60
DB 217 GTGGTTTTTGAATACTTCTGGAGAGACTGTGGATGAATTCGTGGAGCGCTAACATC 276

QY 61 ArgMetGlnCysLysIleTyrAspSerLeuAlaLeuSerProAspLeuGlnAlaAla 80
DB 277 AGAGTGCAGTGCATAATCTATGATTCCTCGTGTGCTTCTTGGCTTTCATGATGGCCATCCT 336

QY 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
DB 337 AGAGGACTGATGTGCTGCTTCCTCGTGTGCTTCTTGGCTTTCATGATGGCCATCCT 396

QY 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
DB 397 GGCATGAATGCACAGGTGCACGGGGACCAATGAGAGGTGAAGGCTCACATTCCTCTG 456

QY 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
DB 457 ACGGCTGGATCATCTTATCAATCAGGGCATGGTGGTGTCTTCTTGGCTTTCATGATGGCCATCCT 516

QY 141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
DB 517 GCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTGGCAAAACGTGAG 576

QY 161 LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla 180
DB 577 CTGGAGAGCTCTCTACTTAGGATGGACACCGCAGCTGGTGTGCTGATTGTGGAGGAGCT 636

QY 181 LeuPheCysValPheCysAsnGluLysSerSerTyrArgTyrSerIlePro 200
DB 637 CTGTTCTGCTGGTTTTTGTGCAACGAAAGAGCAGTACAGTACAGATACCTGATACCT 696

QY 201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSerProSerValTyrSer 220
DB 697 TCCCATCGCACCAACCAAAAGTTATCACACCGGAAAGAGTACCAGGCGTCTACTCC 756

QY 221 ArgSerGlnTyrVal 225
DB 757 AGAAGTCAGTATGTG 771

RESULT 6
AX597259
LOCUS
DEFINITION
Sequence 327 from Patent WO0078961.
ACCESSION
AX597259
VERSION
AX597259.1 GI:29498418
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS

Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L., Eaton, D.L., Gao, W.Q., Pan, J.J., Botstein, D., Fong, S., Goddard, A., Godowski, P.J., Gurney, A.L., Smith, V., Tamas, D., Wood, W.I., Grimaldi, C.J., Hillan, K.J., Paoni, N.F., Roy, M.A. and Watanabe, C.K. Secreted and transmembrane polypeptides and nucleic acids encoding the same

JOURNAL

Patent: WO 0078961-A 327 28-DEC-2000;

FEATURES

Genentech Inc. (US)
Location/Qualifiers
1..2010
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 6,83e-123 Length: 2010
Score: 1172.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-732-120 (1-225) x AX697259 (1-2010)

QY 1 MetAlaThrHisAlaLeuGluLeuAlaGlyLeuPheLeuGlyValGlyMetValGly 20
DB 97 ATGCAACCATGCTTAGAAATCGCTGGCTGTTCTTGGTGTGTTGGAATGTGGGC 156
QY 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnIle 40
DB 157 ACAGTGGCTGTCACTGTCTGCTCAGTGGAGAGTGTGGCCCTTCATTGAAACACATC 216
QY 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle 60
DB 217 GTGGTTTTTGAAACTCTCGGAAGGACTGTGAGTGAATGCGTGGAGCAGCTAACATC 276
QY 61 ArgMetGlnCysLysIleTyAspSerLeuAlaLeuSerProAspLeuGlnAlaLa 80
DB 277 AGGATGCAGTCAAAATCTATGATTCCTGCTGCTCTTCTCCGGACCTACAGGCAGCC 336
QY 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
DB 337 AGAGACGTGATGTGCTGCTCTCGTGATGCTCTTCTGGCTTTCATGATGGCATCCTT 396
QY 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
DB 397 GGCATGAATGACACAGGTGCACGGGGACAAATGAGAAGTGAAGGCTCACATCTCGCTG 456
QY 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
DB 457 ACGGTGGAATCATCTTCATCATCACGGGATGCTGCTCATCTCCCTGTGAGCTGGGT 516
QY 141 AlaAsnAlaIleIleArgAspPheTyAsnSerIleValAsnValAlaGlnLysArgGlu 160
DB 517 GCCAATGCATCATCAGAGATTTCTATACTCAATAGTGAATGTTGCCCAAAACGTGAG 576
QY 161 LeuGlyGluAlaLeuTyLeuGlyTrpThrThrAlaLeuValLeuIleValGlyGlyAla 180
DB 577 CTTCGAGAGAGCTCTCTACTAGGATGGACCAACGCACTGGTGTGCTGATGTTGAGAGCT 636
QY 181 LeuPheCysCysValPheCysCysAsnGluLysSerSerTyTrpArgTySerIlePro 200
DB 637 CTGTTCTCTCTCGGTTTTTTTGTTCACAGAAAGCAGTACGTACAGATCTCGATCCT 696
QY 201 SerHisArgThrGlnLysSerTyHisThrGlyLysLysSerProSerValTySer 220
DB 697 TCCATCGCACAAACCAAAAGTTATCACACCGGAAAGAGTCAACCGAGCGTCTACTCC 756
QY 221 ArgSerGlnTyTrpVal 225

Db 757 AGAAGTCAGTATGTG 771

RESULT 7

AY358707

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-063-732-120 (1-225) x AY358707 (1-2010)

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

AY358707

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-063-732-120 (1-225) x AY358707 (1-2010)

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

RESULT	8
AC131927/c	
LOCUS	HTG 30-DEC-2002
DEFINITION	Homo sapiens chromosome 8 clone RP11-945D10 map 8, LOW-PASS SEQUENCE SAMPLING.
VERSION	AC131927
ACCSSION	AC131927.2 GI:27414257
KEYWORDS	HTG; HTGS PHASE3.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (Bases 1 to 81583) Birren,B., Nusbaum,C. and Lander,E. Homo sapiens chromosome 8, clone RP11-945D10 Unpublished 2 (Bases 1 to 81583) Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Soukngalter,B., Canarata,J., Chang,J.J., Chazarro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menais,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,D., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,

 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L28078
 Center clone name: 945 D 10

*	1	962:	contig of 962 bp	in length
*	963	1062:	gap of 100 bp	
*	1063	1998:	contig of 936 bp	in length
*	1999	2098:	gap of 100 bp	
*	2099	3025:	contig of 927 bp	in length
*	3026	3125:	gap of 100 bp	
*	3126	4085:	contig of 960 bp	in length
*	4086	4185:	gap of 100 bp	
*	4186	5143:	contig of 958 bp	in length
*	5144	5243:	gap of 100 bp	
*	5244	6191:	contig of 948 bp	in length
*	6192	6291:	gap of 100 bp	
*	6292	7252:	contig of 961 bp	in length
*	7253	7352:	gap of 100 bp	
*	7353	8293:	contig of 941 bp	in length
*	8294	8393:	gap of 100 bp	
*	8394	9328:	contig of 935 bp	in length
*	9329	9428:	gap of 100 bp	
*	9429	10373:	contig of 945 bp	in length
*	10374	10473:	gap of 100 bp	
*	10474	11394:	contig of 921 bp	in length

*	11395	11494:	gap of 100 bp	*	49257	50228:	contig of 972 bp in length
*	11495	12397:	contig of 903 bp in length	*	50229	50328:	gap of 100 bp
*	12398	12497:	gap of 100 bp	*	50329	51306:	contig of 978 bp in length
*	12498	13386:	contig of 889 bp in length	*	51307	51406:	gap of 100 bp
*	13387	13486:	gap of 100 bp	*	51407	52385:	contig of 979 bp in length
*	13487	14417:	contig of 931 bp in length	*	52386	52485:	gap of 100 bp
*	14418	14517:	gap of 100 bp	*	52486	53438:	contig of 953 bp in length
*	14518	15479:	contig of 962 bp in length	*	53439	53538:	gap of 100 bp
*	15480	15579:	gap of 100 bp	*	53539	54500:	contig of 962 bp in length
*	15580	16546:	contig of 967 bp in length	*	54501	54600:	gap of 100 bp
*	16547	16647:	gap of 100 bp	*	54601	55512:	contig of 912 bp in length
*	16647	17599:	contig of 953 bp in length	*	55513	55612:	gap of 100 bp
*	17600	17699:	gap of 100 bp	*	55613	56579:	contig of 967 bp in length
*	17700	18667:	contig of 968 bp in length	*	56580	56679:	gap of 100 bp
*	18668	18767:	gap of 100 bp	*	56680	57644:	contig of 965 bp in length
*	18768	19733:	contig of 966 bp in length	*	57645	57744:	gap of 100 bp
*	19734	19833:	gap of 100 bp	*	57745	58703:	contig of 959 bp in length
*	19834	20785:	contig of 952 bp in length	*	58704	58803:	gap of 100 bp
*	20786	20885:	gap of 100 bp	*	58804	59756:	contig of 953 bp in length
*	20886	21824:	contig of 939 bp in length	*	59757	59856:	gap of 100 bp
*	21825	21924:	gap of 100 bp	*	59857	60781:	contig of 925 bp in length
*	21925	22867:	contig of 943 bp in length	*	60782	60881:	gap of 100 bp
*	22868	22967:	gap of 100 bp	*	60882	61804:	contig of 923 bp in length
*	22968	23883:	contig of 916 bp in length	*	61805	61904:	gap of 100 bp
*	23884	23983:	gap of 100 bp				
*	23984	24903:	contig of 920 bp in length				
*	24904	25003:	gap of 100 bp				
*	25004	25971:	contig of 968 bp in length				
*	25972	26071:	gap of 100 bp				
*	26072	27042:	contig of 971 bp in length				
*	27043	27142:	gap of 100 bp				
*	27143	28102:	contig of 960 bp in length				
*	28103	28202:	gap of 100 bp				
*	28203	29173:	contig of 971 bp in length				
*	29174	29273:	gap of 100 bp				
*	29274	30241:	contig of 968 bp in length				
*	30242	30341:	gap of 100 bp				
*	30342	31301:	contig of 960 bp in length				
*	31302	31401:	gap of 100 bp				
*	31402	32340:	contig of 939 bp in length				
*	32341	32440:	gap of 100 bp				
*	32441	33375:	contig of 935 bp in length				
*	33376	33475:	gap of 100 bp				
*	33476	34422:	contig of 947 bp in length				
*	34423	35422:	gap of 100 bp				
*	34523	35449:	contig of 927 bp in length				
*	35450	35549:	gap of 100 bp				
*	35550	36516:	contig of 967 bp in length				
*	36517	36616:	gap of 100 bp				
*	36617	37574:	contig of 958 bp in length				
*	37575	37674:	gap of 100 bp				
*	37675	38660:	contig of 986 bp in length				
*	38661	38760:	gap of 100 bp				
*	38761	39733:	contig of 973 bp in length				
*	39734	39833:	gap of 100 bp				
*	39834	40803:	contig of 970 bp in length				
*	40804	40903:	gap of 100 bp				
*	40904	41870:	contig of 967 bp in length				
*	41871	41970:	gap of 100 bp				
*	41971	42938:	contig of 968 bp in length				
*	42939	43038:	gap of 100 bp				
*	43039	43967:	contig of 929 bp in length				
*	43968	44067:	gap of 100 bp				
*	44068	45006:	contig of 939 bp in length				
*	45007	46023:	contig of 917 bp in length				
*	46024	46123:	gap of 100 bp				
*	46124	47026:	contig of 903 bp in length				
*	47027	47126:	gap of 100 bp				
*	47127	48085:	contig of 959 bp in length				
*	48086	48185:	gap of 100 bp				
*	48186	49156:	contig of 971 bp in length				
*	49157	49256:	gap of 100 bp				

Alignment Scores:

Pred. No.: 7,32e-121 Length: 81583
 Score: 1172.00 Matches: 225
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-063-732-120 (1-225) x AC131927 (1-81583)

Qy	1	MetAlaThrHisAlaLeuGlutAlaGlyLeuPheLeuGlyValGlyMetValGly	20
Db	36440	ATGCCAACCCATGCTAGAAATCGCTGGCTGTTCTTGTGGTGGTGGTGGTGGC	36381
Qy	21	ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnIle	40
Db	36380	ACAGTGGCTGTCACTGTCTCATGCTCAGTGGAGAGTGTGGCTTCATTGAAACAAATC	36321
Qy	41	ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle	60
Db	36320	GTGTTTTTGAACAACTTCGGNAGACGTGTGGTGAATTTGGTGGCAGGCTAACATC	36261
Qy	61	ArgMetGlnCysLysIleTyrAspSerLeuAlaLeuSerProAspLeuGlnAlaIle	80
Db	36260	AGGATGCAGTGCACAAATCTATGATTCCTCTGGTGGTCTTTTCCTCGACCTACAGCAGCC	36201
Qy	81	ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetAlaIleLeu	100
Db	36200	AGAGGACTGAATGTGTCTTCCTGGTGAATGCTTCTTGGCTTTTATGATGGCCATCCTT	36141
Qy	101	GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu	120
Db	36140	GGCATGAAATGCACAGGTGCACGGGGGACAAATGAGAGGTGAGGCTCACATTCTGCTG	36081
Qy	121	ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal	140
Db	36080	ACGGCTGGAATCATCTTCATCATCAGCGGCAGTGGTGGTCTCATCCCTGTAGCTGGGTT	36021
Qy	141	AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu	160
Db	36020	GCCAATGCCATCATCAGAGATTTCATAACTCAATAGTGAATGTTGCCAAAAACGTGAG	35961
Qy	161	LeuGlyGluAlaLeuTyrIleGlyTrpThrThrAlaLeuValLeuIleValGlyGlyAla	180
Db	35960	CTTGGAGAACTCTCTACTTAGGATGGACACGACGCTGGTGGTCTGATTGTGGAGAGCT	35901
Qy	181	LeuPheCysCysValPheCysCysAsnGluLysSerSerTyrArgTyrSerIlePro	200

```

Db      35900 CTGTTCTGCTGCTGTTTGTGTCACCAAGAAAGACGACGTAGCTACAGATCTCGATACCT 35841
QY      201 SerHisArgThrThrGlnLysSerTyrHisThrGlnLysLysSerProSerValTyrSer 220
Db      35840 TCCATCGCACCAACCAAAAGATTATCACCCGGAAGAAGTACCAGGCTCTACTCC 35781
QY      221 ArgSerGlnTyrVal 225
Db      35780 AGAAGTCAGATGTG 35766

RESULT 9
AP000884/c
LOCUS      137955 bp DNA linear PRI 16-DEC-1999
DEFINITION Homo sapiens genomic DNA, chromosome 21q22.1, clone:B463J19,
            SOD-AML region, complete sequence.
ACCESSION AP000884
VERSION    AP000884.1 GI:6580113
KEYWORDS   HTG.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens (human)
REFERENCE  1 (bases 1 to 137955)
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS    Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Submitted (15-DEC-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp).
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924
REFERENCE  2 (bases 1 to 137955)
            Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
TITLE      Submitted (15-DEC-1999) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
            Japan (E-mail:hattori@gsc.riken.go.jp).
            URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
            Fax:81-42-778-9924
FEATURES   Location/Qualifiers
            source          1. .137955
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosomes="21"
                        /map="21q22.1"
                        /clone="B463J19"
ORIGIN
Alignment Scores:
Pred. No.:      1.42e-120      Length:      137955
Score:          1172.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9      Gaps:      0

US-10-063-732-120 (1-225) x AP000884 (1-137955)

QY      1 MetAlaThrHisAlaLeuGlnLysValGlyLeuPheLeuGlyValGlyMetValGly 20
Db      58261 ATGCAACCCATCCCTAGAAATCGTGGCGCTGTTCTTGGGTGGTGGATGGTGGC 58202
QY      21 ThrValAlaValThrValMetProGlnTyrArgValSerAlaPheIleGluAsnAsnIle 40
Db      58201 ACAGTGGCTGCTACTGTCATGCTCAGTGGAGAGTGTCCGCTTCATTGAAACAACATC 58142
QY      41 ValValPheGluAsnPheThrGlnGlyLeuTyrMetAsnCysValArgGlnAlaAsnIle 60
Db      58141 GTGGTGTTCGAAACCTCTCGGAAGACCTGTGGATGAATGGCTGAGGAGGCTTAACATC 58082
QY      61 ArgMetGlnCysLysIleTyrAspSerLeuLeuAlaLeuSerProAspLeuGlnAlaIa 80
Db      58081 AGGATGCAGTGCAAAATCTAATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGCAGCC 58022

```

```

QY      81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
Db      58021 AGAGACTGATGTGTGCTGCTCCGTCATGTCCTCTTGGCTTTCATGATGGCCATCCTT 57962
QY      101 GlyMetLysCysThrArgCysThrGlyAspAsnGlnLysValLysAlaHisIleLeuLeu 120
Db      57961 GGCATGAATGCACAGGTGCACGGGGACAAATGAGAAGGTGAAGGCTCACAATCTCTCTG 57902
QY      121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTyrVal 140
Db      57901 ACGGCTGGAATCATCTTCATCATCACGGGCATGTGGTGTCTCATCCCTGTGAGTGGGTT 57842
QY      141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
Db      57841 GCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTAATGTGCCCAAAACGAGAG 57782
QY      161 LeuGlyGluAlaLeuTyrLeuGlyTyrThrThrAlaLeuValLeuIleValGlyGlyAla 180
Db      57781 CTTGGAGAGCTCTCTACTTAGGATGGACACCGACCTGGTGTGCTGATTGTGGAGGAGCT 57722
QY      181 LeuPheCysCysValPheCysAsnGlnLysSerSerTyrArgTyrSerIlePro 200
Db      57721 CTGTTCTGCTGGTGTGTTTGTTCGAACGAAAGACGACGTAGCTACAGATACCTCGATACCT 57662
QY      201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSerProSerValTyrSer 220
Db      57661 TCCCATCGCACCAACCAAAAGTTATCACCCGGAAGAAGTACCAGGCGCTCTATCC 57602
QY      221 ArgSerGlnTyrVal 225
Db      57601 AGAAGTCAGATGTG 57587

RESULT 10
AP001707/c
LOCUS      340000 bp DNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens genomic DNA, chromosome 21q, section 51/105.
ACCESSION AP001707 AL163252 BA000005
VERSION    AP001707.1 GI:7768786
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
            Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordtsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
            Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
            Yaspo,M.L.
            The DNA sequence of human chromosome 21
            Nature 405 (6784), 311-319 (2000)
            20289799
            MEDLINE
            PUBMED 10830953
            2 (bases 1 to 340000)
            Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
            Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
            Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
            Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
            Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
            Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
            Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
            Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordtsiek,G.,
            Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A.,
            Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K.,
            Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K.,

```



```

/note="(A)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(10285..10467)
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
10468..10753
/note="AluJb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
10757..10785
/note="(AAAA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM

```

Alignment Scores:		
Pred. No.:	4,448-120	Length:
Score:	172.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	100.00%	Indels:
DB:	9	Gaps:
		340000
		225

US-10-063-732-120 (1-225) x AP001707 (1-340000)

Qy	1	MetAlaThrHisAlaLeuGluIleAlaGlyLeuPheLeuGlyValGlyMetValGly	20
Db	214950	ATGGCAACCATCGCTTGAATCGCTGGGCTGTTCTTGGTGGTGTGAATGGTGGGC	214991
Qy	21	ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnAsnIle	40
Db	214890	ACAGTGGCTGTCACCTGTCATGCCTCAGTGGAGAGTGTGGCCCTCATTTGAAAACAACATC	214931
Qy	41	ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle	60
Db	214830	GTGGTTTTTGAATACTTCGGGAAGCACTGTGGATGAATTCGCTGAGCAGGCTAAACATC	214771
Qy	61	ArgMetGlnCysLysIleTyrAsnSerLeuLeuAlaLeuSerProAspLeuGlnAlaAla	80
Db	214770	AGATGTCAGTGCARAAATCATGATTCCTCGTGGCTCTTTCTCCGACCTTACAGGACGC	214711
Qy	81	ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu	100
Db	214710	AGAGGACTCATGTGCTGCTTCCGTGATGCTCTTCTGGCTTTCATGATGGCCATCCTT	214651
Qy	101	GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValIysAlaHisIleLeuLeu	120
Db	214650	GGCATGAATGCACACAGTGCACGGGGACAACTGAGAAGGTGAAGGCTCACATCTGCTG	214591
Qy	121	ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal	140
Db	214590	ACGGCTGGAATCATCTTCATCATCACCGGCGCATGTTGGTGCTCATCCCTGTGAGCTGGTT	214531
Qy	141	AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu	160
Db	214530	GCCAAATGCCATCATCAGAGATTCTATTAACCAATAGTAGAATGTGCCAAAAACGTGAG	214471
Qy	161	LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla	180
Db	214470	CITGGAGAGCTCTACTTAGATGGACACCGCACTGGTGGTGTGATTTGTGGAGGAGCT	214411
Qy	181	LeuPheCysCysValPheCysCysAsnGluLysSerSerTyrArgTyrSerIlePro	200
Db	214410	CTGTTCTGCTGGCTTTTTTTGTTGCAACGAAAGACAGTAGCTACAGATACTCGATACCT	214351
Qy	201	SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSerProSerValTyrSer	220
Db	214350	TCCCATCGACACACCCAAAAAGTTATCACACCGGAAGAAGTACCCGAGCGTCTACTCC	214291
Qy	221	ArgSerGlnTyrVal	225
Db	214290	AGAAGTCAGTATGTG	214276

RESULT 11	
BS00178/c	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

COMMENT

BS000178 176580 bp DNA linear PRI 07-OCT-2000
pan troglodytes chromosome 22 clone:RP43-042C06, map 22, complete
sequences.
BS000178
BS000178.1 GI:37537445
HTG
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
The Chimpanzee Chromosome 22 Sequencing Consortium.
DNA sequence of chimpanzee chromosome 22 and its evolutionary
implications
Unpublished
2 (bases 1 to 176580)
Wang,S., Cai,Z., Wang,B., Zheng,H., Zhang,Y., Zhang,X., Zhu,G.,
Lu,G., Fu,G. and Chen,Z.
Direct Submission
Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human
Genome Center at Shanghai, Genomic Sequencing: No.250 BiBo Road,
Zhang Jiang Hi-TECH Park, Shanghai 201203, CHINA
(E-mail:wangsc@chgc.sh.cn, URL:http://www.chgc.sh.cn,
Tel:86-21-50801919, Fax:86-21-50801922)
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*GDF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRISS Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.
----- Genome Center
Center: Chinese National Human Genome Center at Shanghai Center
code: CHGCS
Web site: http://chgc.sh.cn
Contact: wangsc@chgc.sh.cn
----- Project Information
Center project name:the Chimpanzee Chromosome 22 Sequencing Project
Center clone name: RP43-042C06
----- Summary Statistics
Sequencing vector: pUC18,100% of reads
Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly
Program: Phrap; version 0.950329
Consensus quality: 176318 bases at least Q40
Consensus quality: 176550 bases at least Q30
Consensus quality: 176577 bases at least Q20
Quality coverage: 9.0x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.

Source information:
The RP43 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pBAC3.6
Sequence Quality Assessment:
This entry has been annotated with sequence

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: PTB-118H03 (left) and RP43-082J09 (right).

FEATURES

source
1..176580
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="RP43-042C06"
/clone_lib="RP43 chimpanzee BAC"

ORIGIN

Alignment Scores:	7,17e-120	Length:	176580
Prod. No.:	1167.00	Matches:	224
Percent Similarity:	99.56%	Conservative:	0
Best Local Similarity:	99.56%	Mismatches:	1
Query Match:	99.57%	Indels:	0
DB:	9	Gaps:	0

US-10-063-732-120 (1-225) x BS000178 (1-176580)

QY	1	MetAlaThrHisAlaLeuGluIleAlaGlyLeuPheLeuGlyValGlyMetValGly	20
Db	31108	ATGGCAACCCAGCGCTTAGAATCGCTGGCTGTTCTTGCTGGTGTGGAAATGGTGGCG	31049
QY	21	ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnIle	40
Db	31048	ACAGTGGCTGTCACCTGTCATCCCTCAGTGGAGAGTGTGGCCCTCATTGAAACACATC	30989
QY	41	ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnIle	60
Db	30988	GTGGTTTTTGAACACTTCTGGGAAGGACTGTGGATGAATTCGTGGAGCGGCTAACATC	30929
QY	51	ArgMetGlnCysLysIleTyrAspSerLeuLeuAlaLeuSerProAspLeuGlnAlaIle	80
Db	30928	AGGATGCAGTGCACAAATCTATGATTCCTCTGCTGCTCTTCTCCGACCTACAGGCGAGC	30869
QY	81	ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu	100
Db	30868	AGAGGACTGATGTGCTGCTTCGTGATGTCCTCTCTTCTGCTTCATGATGGCCATCTT	30809
QY	101	GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu	120
Db	30808	GGCATGAAATGCACAGGTGCACGGGGGCAATGAGAAGGTGAAGGCTCACATTTCTGCTG	30749
QY	121	ThrAlaGlyIleIlePheIleThrGlyMetValValLeuIleProValSerTrpVal	140
Db	30748	ACGGCTGGANCTCTTCATCATCAGCGGCAATGCTGGTGTCTCATCTCTGTGAGCTGGTT	30689
QY	141	AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu	160
Db	30688	GCCAAATGCCATCATCAGAGATTTCTATAACCCCAATAGTGAATGTGCCAAAAACGTGAG	30629
QY	161	LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla	180
Db	30628	CTTGAGAGCTCTCTACTTAGGATGACCCAGCGCTGTGCTGATTTGTGGAGAGCT	30569
QY	181	LeuPheCysCysValPheCysCysAsnGluLysSerSerTyrArgTyrSerIlePro	200
Db	30568	CTGTTCTGCTCGTTTTTTTGTTCACAGAAAAGAGCAGTAGCTACAGATACGATACCT	30509
QY	201	SerHisArgThrGlnLysSerTyrHisThrGlyLysSerProSerValTyrSer	220
Db	30508	TCCCATGCAACACCAAAAGATTATCACCGGAAAGAGTACCCGAGCGCTCTACTCC	30449
QY	221	ArgSerGlnTyrVal	225
Db	30448	AGAAGTCAGTATGTG	30434

RESULT 12

BS000177/c

LOCUS

BS000177.1

DEFINITION

Pan troglodytes chromosome 22 clone:PTB-118H03, map 22, complete

ACCESSION

BS000177

VERSION

BS000177.1

KEYWORDS

HTG.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

REFERENCE

1

AUTHORS

Wang,S., Cai,Z., Wang,B., Zheng,H., Zhang,Y., Zhang,X., Zhu,G.,

Lu,G., Fu,G. and Chen,Z.

TITLE

Direct Submission

JOURNAL

Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human

Genome Center at Shanghai, Genomic Sequencing; No.250 BiBo Road,

Zhang Jiang HI-TECH Park, Shanghai 201203, CHINA

(E-mail:wangsy@chgc.sh.cn, URL:http://www.chgc.sh.cn,

Tel:86-21-50801919, Fax:86-21-50801922)

The Chimpanzee Chromosome 22 Sequencing Consortium consists of:

*Chinese National Human Genome Center at Shanghai, Shanghai, China;

*GSF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute

of Molecular Biotechnology, Jena, Germany; *KRIIBB Genome Research

Center, Daejeon, Korea;

*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;

*National Institute of Genetics, Mishima, Japan;

*National Yang Ming University Genome Research Center, Taipei,

Taiwan;

*RIKEN Genomic Sciences Center, Yokohama, Japan.

Center: Chinese National Human Genome Center at Shanghai Center

code: CHGC

Web site: http://chgc.sh.cn

Contact: wangsy@chgc.sh.cn

Project Information

Center project name: The Chimpanzee Chromosome 22 Sequencing Project

Center clone name: PTB-118H03

Summary Statistics

Sequencing vector: pUC18/100% of reads

Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly

program: Phrap; version 0.990429

Consensus quality: 266876 bases at least Q40

Consensus quality: 267150 bases at least Q30

Consensus quality: 267168 bases at least Q20

Quality coverage: 9.3x

This sequence was finished as follows unless otherwise noted: all

regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30);

an attempt was made to resolve all sequencing problems, such as

compressions and repeats; all regions were covered by at one

subclone or more than one M13 subclone;

and the assembly was confirmed by restriction digest.

Source information:

The PTB1 chimpanzee BAC library was prepared from DNA isolated from

cultured cells established from the blood of a single male

chimpanzee.

Clones may be obtained from Asao Fujiyama and co-workers

(http://www.gsc.riken.go.jp).

VECTOR: pKS145

Sequence Quality Assessment:

This entry has been annotated with sequence

BS000177 267172 bp DNA linear PRI 07-OCT-2003
Pan troglodytes chromosome 22 clone:PTB-118H03, map 22, complete
sequences.

BS000177 GI:37537444

HTG.

Pan troglodytes (chimpanzee)

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

The Chimpanzee Chromosome 22 Sequencing Consortium.

DNA sequence of chimpanzee chromosome 22 and its evolutionary

implications

Unpublished

2 (bases 1 to 267172)

Wang,S., Cai,Z., Wang,B., Zheng,H., Zhang,Y., Zhang,X., Zhu,G.,

Lu,G., Fu,G. and Chen,Z.

Direct Submission

Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human

Genome Center at Shanghai, Genomic Sequencing; No.250 BiBo Road,

Zhang Jiang HI-TECH Park, Shanghai 201203, CHINA

(E-mail:wangsy@chgc.sh.cn, URL:http://www.chgc.sh.cn,

Tel:86-21-50801919, Fax:86-21-50801922)

The Chimpanzee Chromosome 22 Sequencing Consortium consists of:

*Chinese National Human Genome Center at Shanghai, Shanghai, China;

*GSF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute

of Molecular Biotechnology, Jena, Germany; *KRIIBB Genome Research

Center, Daejeon, Korea;

*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;

*National Institute of Genetics, Mishima, Japan;

*National Yang Ming University Genome Research Center, Taipei,

Taiwan;

*RIKEN Genomic Sciences Center, Yokohama, Japan.

Center: Chinese National Human Genome Center at Shanghai Center

code: CHGC

Web site: http://chgc.sh.cn

Contact: wangsy@chgc.sh.cn

Project Information

Center project name: The Chimpanzee Chromosome 22 Sequencing Project

Center clone name: PTB-118H03

Summary Statistics

Sequencing vector: pUC18/100% of reads

Chemistry: Dye-terminator Big Dye and ET; 100% of reads Assembly

program: Phrap; version 0.990429

Consensus quality: 266876 bases at least Q40

Consensus quality: 267150 bases at least Q30

Consensus quality: 267168 bases at least Q20

Quality coverage: 9.3x

This sequence was finished as follows unless otherwise noted: all

regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30);

an attempt was made to resolve all sequencing problems, such as

compressions and repeats; all regions were covered by at one

subclone or more than one M13 subclone;

and the assembly was confirmed by restriction digest.

Source information:

The PTB1 chimpanzee BAC library was prepared from DNA isolated from

cultured cells established from the blood of a single male

chimpanzee.

Clones may be obtained from Asao Fujiyama and co-workers

(http://www.gsc.riken.go.jp).

VECTOR: pKS145

Sequence Quality Assessment:

This entry has been annotated with sequence

estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: RP43-006021 (left) and RP43-042C06 (right).

FEATURES

source
Location/Qualifiers
1..267172
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="PTB-118H03"
/clone_lib="PTB1 chimpanzee BAC"

ORIGIN

Alignment Scores:
Pred. No.: 1,21e-119 Length: 267172
Score: 1167.00 Matches: 224
Percent Similarity: 99.56% Conservative: 0
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.57% Indels: 0
DB: 9 Gaps: 0

US-10-063-732-120 (1-225) x BS000177 (1-267172)

QY	1	MetAlaThrHisAlaLeuGlulLeuAlaGlyLeuPheLeuGlyGlyValGlyMetValGly	20
DB	258767	ATGGCAACCCATGCTTGAATCGCTGGCTGTTCTTGGTGTGTGAATGGTGGC	258708
QY	21	ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheLeuGluAsnAsnile	40
DB	258707	ACAGTGGCTGTACATGTCATGCTCAGTGGAGAGTGTGGCTTCAITGAAACCAACATC	258648
QY	41	ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnile	60
DB	258647	GTGGTGTGTGAATCTCTGGGAAGGACTGTGGATGAATTCGTGGAGCGAGCTTAACATC	258588
QY	61	ArgMetGlnCysLysLeuTrpAspSerLeuLeuAlaLeuSerProAspLeuGlnAlaAla	80
DB	258587	AGATGCACTGCAAAATCTATGATTCCTGTGGCTCTTTCTCGGACCTACAGGCGACC	258528
QY	81	ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaileLeu	100
DB	258527	AGAGGACTGATGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTT	258468
QY	101	GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisileLeuLeu	120
DB	258467	GGCATGAAATGCCAGGTGCACGGGGACATGATGAAGGTGAAGGCTCACATTTCTGCTG	258408
QY	121	ThrAlaGlyIlePheLeuLeuThrGlyMetValValLeuLeuLeuProValSerTrpVal	140
DB	258407	ACGGCTGGAATCATCTTCATCATCAGGGCATGGTGGTGTCTATCCCTGTGAGCTGGTT	258348
QY	141	AlaAsnAlaIlelleArgAspPheTrpAsnSerIleValAsnValAlaGlnLysArgGlu	160
DB	258347	GCCAAATGCCATCATCAGAGATTTCTATAACCAATAGTGAATGTGCCCAAAACGFGAG	258288
QY	161	LeuGlyGluAlaLeuTrpLeuGlyTrpThrAlaLeuValLeuLeuValcylGlyAla	180
DB	258287	CTTGGAGAGCTCTTACTTAGATGGACCGGACCTGGTGTGTGATTTGTGGAGAGCT	258228
QY	181	LeuPheCysValPheCysAsnGluLysSerSerTrpArgTrpSerIlePro	200
DB	258227	CTGTTCTGCTGGCTTTTGTGTGCAACGAAAGAGCAGTAGCTACAGATCTCGATACCT	258168
QY	201	SerHisArgThrGlnLysSerTrpHisThrGlyLysLysSerProSerValTrpSer	220
DB	258167	TCCATCGCACCAACCCAAAGAGTTATCACCGGAAAGAGTACCGAGGCTCTACTCC	258108
QY	221	ArgSerGlnTrpVal	225
DB	258107	AGAAGTCAGTATGTG	258093

RESULT 13

AX866683
LOCUS
Sequence 1588 from Patent EP1074617.
DEFINITION
AX866683
VERSION
AX866683.1 GI:40021032
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primers for synthesising full-length cDNA and their use
Patent: EP 1074617-A 1588 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9,72e-120 Length: 837
Score: 1140.00 Matches: 221
Percent Similarity: 98.23% Conservative: 1
Best Local Similarity: 97.79% Mismatches: 3
Query Match: 97.27% Indels: 1
DB: 6 Gaps: 0

US-10-063-732-120 (1-225) x AX866683 (1-837)

QY	1	MetAlaThrHisAlaLeuGlulLeuAlaGlyLeuPheLeuGlyGlyValGlyMetValGly	20
DB	150	ATGGCAACCCATGCTTGAATCGCTGGCTGTTCTTGGTGTGTGAATGGTGGC	209
QY	21	ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheLeuGluAsnAsnile	40
DB	210	ACAGTGGCTGTACATGTCATGCTCAGTGGAGAGTGTGGCTTCAITGAAACCAACATC	269
QY	41	ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaAsnile	60
DB	270	GTGGTGTGTGAATCTCTGGGAAGGACTGTGGATGAATTCGTGGAGCGAGCTTAACATC	329
QY	61	ArgMetGlnCysLysLeuTrpAspSerLeuLeuAlaLeuSerProAspLeuGlnAlaAla	80
DB	330	AGATGCACTGCAAAATCTATGATTCCTGTGGCTCTTTCTCGGACCTACAGGCGACC	389
QY	81	ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaileLeu	100
DB	390	AGAGGACTGATGTGCTGCTTCCGTGATGCCCTTCTTGGCTTTCATGATGGCCATCCTT	449
QY	101	GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisileLeuLeu	120
DB	450	GGCATGAAATGCCAGGTGCACGGGGACATGATGAAGGTGAAGGCTCACATTTCTGCTG	509
QY	121	ThrAlaGlyIlellePheLeuLeuThrGlyMetValValLeuLeuLeuProValSerTrpVal	140
DB	510	ACGGCTGGAATCATCTTCATCATCAGGGCATGGTGGTGTCTATCCCTGTGAGCTGGTT	569
QY	141	AlaAsnAlaIlelleArgAspPheTrpAsnSerIleValAsnValAlaGlnLysArgGlu	160
DB	570	GCCAAATGCCATCATCAGAGATTTCTATAACCAATAGTGAATGTGCCCAAAACGFGAG	629
QY	161	LeuGlyGluAlaLeuTrpLeuGlyTrpThrAlaLeuValLeuLeuValcylGlyAla	180
DB	630	CTTGGAGAGCTCTTACTTAGATGGACCGGACCTGGTGTGTGATTTGTGGAGAGCT	689
QY	181	LeuPheCysValPheCysAsnGluLysSerSerTrpArgTrpSerIlePro	200

```

Db      630 CTGTTCTGCTGGCTTTTCTGTCACGAAAGAGCAGTAGCTACAGATACTCGATACCT 749
Qy      201 SerHisArgThrThrGlnLysSerTyHisThrGlyLysLysSer-ProSerValTyrSe 220
Db      750 TCCATCGCACAAACCAAAAGTTATCACCGGAAAGAGTACACCGACCGTNTACTT 809
Qy      220 rArgSerGlnTyrVal 225
Db      810 CAGAAGTCAGTATGTG 825

RESULT 14
BD146745
LOCUS      837 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD146745
VERSION      BD146745.1 GI:27852503
KEYWORDS      JP 2002191363-A/1588.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 837)
Ota,i., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1588 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/1588
PD      09-JUL-2002
PF      28-JUL-2000 JP 2000280990
PI      TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI      SAITO,
PI      JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI      KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC      C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT      source
FT      Location/Qualifiers
1..837
/organism="Homo sapiens (human)"
/locus="BD146745"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      9.72e-120      Length:      837
Score:      1140.00      Matches:      221
Percent Similarity:      98.23%      Conservative:      1
Best Local Similarity:      97.79%      Mismatches:      3
Query Match:      97.27%      Indels:      1
Db:      6      Gaps:      0

US-10-063-732-120 (1-225) x BD146745 (1-837)

Qy      1 MetAlaThrHisAlaLeuGluIleAlaGlyLeuPheLeuGlyValGlyMetValGly 20
Db      150 ATGGCAACCCATGCTTATAGAAATCGCTGGCTGTTCTTCTGGTGTGTAATGTGGGC 209
Qy      21 ThrValAlaValThrValMetProGlnTyrArgValSerAlaPheIleGluAsnIle 40
Db      210 ACATGTGCTGTCATGTGCCTCATGTGAGAGTGTGGCTTCATTTGAAACCAACATC 269
Qy      41 ValValPheGluAsnPheTyrGluGlyLeuTyrMetAsnGlyValArgGlnAlaAsnIle 60
Db      270 GTGGTTTGTGAAACTTCTGGAGAGGACTGTGGATGAATTCGTGAGGAGGCTAACATC 329
Qy      61 ArgMetGlnCysLysIleTyrAspSerLeuLeuAlaLeuSerProAspLeuGlnAla 80

```

```

Db      330 AGATGCAGTGCAGAAATCTATGATTCCTCTGCTGGCTCTTTCTCGGACCTACAGCAGCC 389
Qy      81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetValAlaIleLeu 100
Db      390 AGAGGACTGATGTGTGCTGCTTCCGTGATGCCCTCTTCTGGCTTTTCATGATGCCCATCCTT 449
Qy      101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
Db      450 GGCATGAATGCACAGGTGCACGGGGACATGAGAGGTGAAGGCTCACATTCGTGCTG 509
Qy      121 ThrAlaGlyIleIlePheIleThrGlyMetValValLeuIleProValSerTrpVal 140
Db      510 ACGGTGGAATCATCTTCATCATCAGCGGCGCTGCTCATCCTCTGTGAGCTGGGTT 569
Qy      141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
Db      570 GCCAATGCCATCATCAGAGATTCTATATACCCCAATAGTAGAATGTTGCCCAAAACGTCGAG 629
Qy      161 LeuGlyGluAlaLeuTyrLeuGlyTyrThrThrAlaLeuValLeuIleValGlyGlyAla 180
Db      630 CTTGGAGAGCTCTCTACTTAGGATGGACCGCACCTGGTCTGATTTGGAGGACT 689
Qy      181 LeuPheCysCysValPheCysCysAsnGluLysSerSerTyrArgTyrSerIlePro 200
Db      690 CTGTTCTGCTGGCTTTTGTTCACACGAAAGACAGTAGCTACAGATCTCGATACCT 749
Qy      201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysLysSer-ProSerValTyrSe 220
Db      750 TCCATCGCACAAACCAAAAGTTATCACCGGAAAGAGTACACCGACCGTNTACTT 809
Qy      220 rArgSerGlnTyrVal 225
Db      810 CAGAAGTCAGTATGTG 825

RESULT 15
AP001846
LOCUS      191923 bp      DNA      linear      HTG 30-MAY-2000
DEFINITION      Homo sapiens chromosome 18 clone RP11-796C24 map 18q22, WORKING
DRAFT SEQUENCE, 56 unordered pieces.
ACCESSION      AP001846
VERSION      AP001846.2 GI:8117509
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 191923)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 191,923 genomic DNA of 18q22
Published Only in DataBase (2000)
2 (bases 1 to 191923)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (18-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 30, 2000 this sequence version replaced gi:7630218.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-796C24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads

```

Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 167257 bases at least Q40
Consensus quality: 178235 bases at least Q30
Consensus quality: 183393 bases at least Q20
Insert size: 186423; sum-of-contigs
Quality coverage: 4.34x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 56 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 10920 contig of 10920 bp in length
11021 18635 contig of 7615 bp in length
18736 27618 contig of 8883 bp in length
27719 33370 contig of 5652 bp in length
3471 38522 contig of 5052 bp in length
33471 49335 contig of 5022 bp in length
38623 44936 contig of 5022 bp in length
45036 52903 contig of 5022 bp in length
53004 59933 contig of 6930 bp in length
60034 65715 contig of 6930 bp in length
65816 73615 contig of 7800 bp in length
73716 79180 contig of 5465 bp in length
79281 83364 contig of 4084 bp in length
83465 91199 contig of 4045 bp in length
91300 94966 contig of 3667 bp in length
95067 99407 contig of 4341 bp in length
99508 103203 contig of 3696 bp in length
103304 106888 contig of 3585 bp in length
106989 110187 contig of 3199 bp in length
110288 113798 contig of 3511 bp in length
113899 117663 contig of 3372 bp in length
117764 121135 contig of 3372 bp in length
121236 123224 contig of 2730 bp in length
123225 125704 contig of 2286 bp in length
125705 127991 contig of 2593 bp in length
127991 130683 contig of 2593 bp in length
130683 133711 contig of 1584 bp in length
133711 137988 contig of 1728 bp in length
137988 141388 contig of 1342 bp in length
141388 144889 contig of 1260 bp in length
144889 146786 contig of 1366 bp in length
146786 149573 contig of 1453 bp in length
149573 151996 contig of 1312 bp in length
151996 154941 contig of 491 bp in length
154941 157109 contig of 2441 bp in length
157109 159650 contig of 2441 bp in length
159650 161167 contig of 2176 bp in length
161167 163443 contig of 2176 bp in length
163443 165635 contig of 2092 bp in length
165635 168432 contig of 2697 bp in length
168432 171331 contig of 2799 bp in length
171331 173015 contig of 1584 bp in length
173015 174843 contig of 1728 bp in length
174843 176285 contig of 1342 bp in length
176285 177628 contig of 1243 bp in length
177628 179422 contig of 1694 bp in length
179422 180706 contig of 1184 bp in length
180706 182215 contig of 1409 bp in length
182215 183673 contig of 1358 bp in length
183673 185031 contig of 1260 bp in length
185031 186901 contig of 1768 bp in length
186901 188367 contig of 1366 bp in length
188367 189920 contig of 1453 bp in length
189920 191332 contig of 1312 bp in length
191332 191923 contig of 491 bp in length

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 10920: contig of 10920 bp in length
10921 11020: gap of 100 bp
11021 18635: contig of 7615 bp in length
18636 18735: gap of 100 bp
18736 27618: contig of 8883 bp in length
27619 27718: gap of 100 bp
27719 33370: contig of 5652 bp in length
33371 33470: gap of 100 bp
33471 38522: contig of 5052 bp in length
38523 44935: contig of 5022 bp in length
44936 49335: gap of 100 bp
49336 52903: contig of 5022 bp in length
52904 59933: contig of 6930 bp in length
59934 60033: gap of 100 bp
60034 65715: contig of 6930 bp in length
65716 73615: gap of 100 bp
73616 79180: contig of 5465 bp in length
79181 83364: contig of 4084 bp in length
83365 83464: gap of 100 bp
83465 87054: contig of 3590 bp in length
87055 87154: gap of 100 bp
87155 91199: contig of 4045 bp in length
91200 94966: contig of 3667 bp in length
94967 95066: gap of 100 bp
95067 99407: contig of 4341 bp in length
99408 99507: gap of 100 bp
99508 103203: contig of 3696 bp in length
103204 103303: gap of 100 bp
103304 106888: contig of 3585 bp in length
106889 110187: contig of 3199 bp in length
110188 110287: gap of 100 bp
110288 113798: contig of 3511 bp in length
113799 113898: gap of 100 bp
113899 117663: contig of 3372 bp in length
117664 121135: gap of 100 bp
121136 121235: gap of 100 bp
121236 123224: contig of 1989 bp in length
123225 123224: gap of 100 bp
123225 125704: contig of 2280 bp in length
125705 127990: contig of 2286 bp in length
127991 128090: gap of 100 bp
128091 130683: contig of 2593 bp in length
130684 130783: gap of 100 bp
130784 133711: contig of 2588 bp in length
133712 134721: gap of 100 bp
134722 136426: contig of 2355 bp in length
136427 136526: gap of 100 bp
136527 139256: contig of 2730 bp in length
139257 139356: gap of 100 bp
139357 141388: contig of 2032 bp in length
141389 141488: gap of 100 bp
141489 144089: contig of 2601 bp in length
144090 144189: gap of 100 bp
144190 146686: contig of 2497 bp in length
146687 146786: gap of 100 bp
146787 149573: contig of 3087 bp in length
149574 149973: gap of 100 bp

Qy 221 ArgSerGlnTyrVal 225
Db 130069 AGAAGTCAGTATGTG 130083
Search completed: September 3, 2004, 04:36:28
Job time : 2908 secs

Qy 151895: contig of 1922 bp in length
* 149974 151895: gap of 100 bp
* 151896 154840: contig of 2845 bp in length
* 154841 154940: gap of 100 bp
* 154941 157109: contig of 2169 bp in length
* 157110 157209: gap of 100 bp
* 157210 159650: contig of 2441 bp in length
* 159651 159750: gap of 100 bp
* 159751 161167: contig of 1417 bp in length
* 161168 161267: gap of 100 bp
* 161268 163443: contig of 2176 bp in length
* 163444 163543: gap of 100 bp
* 163544 165635: contig of 2092 bp in length
* 165636 165735: gap of 100 bp
* 165736 168432: contig of 2697 bp in length
* 168433 168532: gap of 100 bp
* 168533 171331: contig of 2799 bp in length
* 171332 171431: gap of 100 bp

Alignment Scores:
Pred. No.: 4.45e-116 Length: 191923
Score: 1134.00 Matches: 222
Percent Similarity: 98.67% Conservative: 0
Best Local Similarity: 98.67% Mismatches: 3
Query Match: 96.76% Indels: 1
DB: 2 Gaps: 0

US-10-063-732-120 (1-225) x AP001846 (1-191923)

Qy 1 MetAlaThrHisAlaLeuGluIleAlaGlyLeuPheLeuGlyValGlyMetValGly 20
Db 129410 ATGGCAACCCATGCTTAAA-ATCGCTGGGCTGTTCCTGGTGTGTGGAAATGGTGGGC 129468

Qy 21 ThrValAlaValThrValMetProGlnTrpArgValSerAlaPheIleGluAsnIle 40
Db 129469 ACAGTGGCTGTCTACCTGTCATGCTCAGTGGAGAGTGTGGCCCTTCATTGAAAAACAACATC 129528

Qy 41 ValValPheGluAsnPheTrpGluGlyLeuTrpMetAsnCysValArgGlnAlaIle 60
Db 129529 GTGGTTTITGAAACCTTCTGGGAAGGACTGTGGATGAATTCGTGGAGGAGGCTAACATC 129588

Qy 61 ArgMetGlnCysLysIleTyrAspSerLeuAlaLeuSerProAspLeuGlnAlaAla 80
Db 129589 AGGATGCAGTGCATAAATCTATGATTCCTGCTGGCTCTTCTCCGGACCTACAGGAGGC 129648

Qy 81 ArgGlyLeuMetCysAlaAlaSerValMetSerPheLeuAlaPheMetMetAlaIleLeu 100
Db 129649 AGAGGACTGATGTGTGCTGCTTCGTGATGCTCCTTCTTGGCTTTTCATGATGGCCATCCTT 129708

Qy 101 GlyMetLysCysThrArgCysThrGlyAspAsnGluLysValLysAlaHisIleLeuLeu 120
Db 129709 GGCAATGAATGCCACAGGTGCACGGGGGACAAATGAGAGGTGAAGGCTCACATTCTGCTG 129768

Qy 121 ThrAlaGlyIleIlePheIleIleThrGlyMetValValLeuIleProValSerTrpVal 140
Db 129769 ACGGCTGGAAATCATCTTCATCATCGCGGCTGGTGGTGTCTATCCCTGTGAGCTGGGTT 129828

Qy 141 AlaAsnAlaIleIleArgAspPheTyrAsnSerIleValAsnValAlaGlnLysArgGlu 160
Db 129829 GCCAATGCCATCATCAGAGATTTCTATACCCCAATAGTGAATGTGCCAAAAACGTGAG 129888

Qy 161 LeuGlyGluAlaLeuTyrLeuGlyTrpThrAlaLeuValLeuIleValGlyGlyAla 180
Db 129889 CTTTGAGAAGCTCTCTACTTAGGATGGACCCAGCCACTGGTGTGATTGTTGGAGAGCT 129948

Qy 181 LeuPheCysCysValPheCysAsnGluLysSerSerSerTyrArgTyrSerIlePro 200
Db 129949 CTGTCTGCTGCTTTTGTGTGACAAAGAGCAGTAGCTACAGATACCTCGATACCT 130008

Qy 201 SerHisArgThrThrGlnLysSerTyrHisThrGlyLysSerProSerValTyrSer 220
Db 130009 TCCCATCGACAAACCCAAAAAAGTTATCACCCGGAAGAGTACCAGAGCGTCTACTCC 130068

Qy 221 ArgSerGlnTyrVal 225
 Db 130069 AGAAGTCAGTATGTG 130083
 Search completed: September 3, 2004, 04:36:28
 Job time : 2908 secs